

This page is not part of
the document!

US2003017559 / 2003-101400

1/2

Date: Dec 11, 2003

Recipient: IB

This Page Blank (uspto)

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
11 December 2003 (11.12.2003)

PCT

(10) International Publication Number
WO 03/101400 A2

- (51) International Patent Classification⁷: **A61K**
- (21) International Application Number: **PCT/US03/17559**
- (22) International Filing Date: **4 June 2003 (04.06.2003)**
- (25) Filing Language: **English**
- (26) Publication Language: **English**
- (30) Priority Data:
60/385,505 **4 June 2002 (04.06.2002)** **US**
- (71) Applicant (*for all designated States except US*): **AVALON PHARMACEUTICALS, INC.** [US/US]; 20358 Seneca Meadows Parkway, Germantown, MD 20876 (US).
- (72) Inventors; and
- (75) Inventors/Applicants (*for US only*): **SHEA, Martin** [US/US]; 345 Market Street West #316, Gaithersburg, MD 20878 (US). **EBNER, Reinhard** [US/US]; 9906 Shelburn Terrace #316, Gaithersburg, MD 20878 (US).
- (74) Agents: **GRANT, Alan, J.** et al.; Carella, Byrne, Bain, Gilfillan, Cecchi, Stewart &, Olstein, 6 Becker Farm Road, Roseland, NJ 07068 (US).

- (81) Designated States (*national*): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NI, NO, NZ, OM, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.
- (84) Designated States (*regional*): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PT, RO, SE, SI, SK, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:

— *without international search report and to be republished upon receipt of that report*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: **CANCER-LINKED GENE AS TARGET FOR CHEMOTHERAPY**

(57) Abstract: Cancer-linked gene sequences, and derived amino acid sequences, are disclosed along with processes for assaying potential antitumor agents based on their modulation of the expression of these cancer-linked genes. Also disclosed are antibodies that react with the disclosed polypeptides and methods of using the antibodies to treat cancerous conditions, such as by using the antibody to target cancerous cells *in vivo* for purposes of delivering therapeutic agents thereto. Also described are methods of diagnosing using the gene sequences.



WO 03/101400 A2

This Page Blank (uspto)

CANCER-LINKED GENE AS TARGET FOR CHEMOTHERAPY

5

This application claims priority of U.S. Provisional Patent Application
60/385,505, filed 4 June 2002, the disclosure of which is hereby incorporated
10 by reference in its entirety.

15

FIELD OF THE INVENTION

The present invention relates to methods of screening cancer-linked
genes and expression products for involvement in the cancer initiation and
20 facilitation process as a means of cancer diagnosis as well as the use of such
genes for screening potential anti-cancer agents, including small organic
compounds and other molecules, and development of therapeutic agents.

25

BACKGROUND OF THE INVENTION

Cancer-linked genes are valuable in that they indicate genetic
differences between cancer cells and normal cells, such as where a gene is
30 expressed in a cancer cell but not in a non-cancer cell, or where said gene is
over-expressed or expressed at a higher level in a cancer as opposed to
normal or non-cancer cell. In addition, the expression of such a gene in a
normal cell but not in a cancer cell, especially of the same type of tissue, can
indicate important functions in the cancerous process. For example, screening
35 assays for novel drugs are based on the response of model cell based

systems *in vitro* to treatment with specific compounds. Such genes are also useful in the diagnosis of cancer and the identification of a cell as cancerous. Gene activity is readily measured by measuring the rate of production of gene products, such as RNAs and polypeptides encoded by such genes. Where
5 genes encode cell surface proteins, appearance of, or alterations in, such proteins, as cell surface markers, are an indication of neoplastic activity. Some such screens rely on specific genes, such as oncogenes (or gene mutations). In accordance with the present invention, a cancer-linked gene has been identified and its putative amino acid sequence worked out. Such
10 gene is useful in the diagnosing of cancer, the screening of anticancer agents and the treatment of cancer using such agents, especially in that these genes encode polypeptides that can act as markers, such as cell surface markers, thereby providing ready targets for anti-tumor agents such as antibodies, preferably antibodies complexed to cytotoxic agents, including apoptotic
15 agents.

BRIEF SUMMARY OF THE INVENTION

20

In accordance with the present invention, there is provided herein a cancer specific gene, linked especially to kidney cancer, or otherwise involved in the cancer initiating and facilitating process and the derived amino acid sequence thereof, including a number of different transcripts derived from said
25 gene.

In one aspect, the present invention relates to a process for identifying an agent that modulates the activity of a cancer-related gene comprising:

(a) contacting a compound with a cell containing a gene that
30 corresponds to a polynucleotide having a sequence selected from the group consisting of SEQ ID NO: 1-7, 14-20 and 27-33, and under conditions promoting the expression of said gene; and

(b) detecting a difference in expression of said gene relative to when said compound is not present

thereby identifying an agent that modulates the activity of a cancer-related gene.

5

In various embodiments of such a process, the cell is a cancer cell and the difference in expression is a decrease in expression. Such polynucleotides may also include those that have sequences identical to SEQ ID NO: 1-7, 14-20 and 27-33.

10

In another aspect, the present invention relates to a process for identifying an anti-neoplastic agent comprising contacting a cell exhibiting neoplastic activity with a compound first identified as a cancer related gene modulator using an assay process disclosed herein and detecting a decrease in said neoplastic activity after said contacting compared to when said contacting does not occur. Such neoplastic activity may include accelerated cellular replication and/or metastasis, and the decrease in neoplastic activity preferably results from the death of the cell, or senescence, terminal differentiation or growth inhibition.

20

The present invention also relates to a process for identifying an anti-neoplastic agent comprising administering to an animal exhibiting a cancer condition an effective amount of an agent first identified according to a process of one of one of the assays disclosed according to the invention and detecting a decrease in said cancerous condition.

The present invention further relates to a process for determining the cancerous status of a cell, comprising determining an increase in the level of expression in said cell of at least one gene that corresponds to a polynucleotide having a sequence selected from the group consisting of SEQ ID NO: 1-7, 14-20 and 27-33 wherein an elevated expression relative to a known non-cancerous cell indicates a cancerous state or potentially

30

cancerous state. Such elevated expression may be due to an increased copy number.

5 The present invention additionally relates to an isolated polypeptide, encoded by one of the polynucleotide transcripts disclosed herein, comprising an amino acid sequence homologous to an amino acid sequence selected from the group consisting of SEQ ID NO: 8-13, 21-26 and 34-39, wherein any difference between said amino acid sequence and the sequence of SEQ ID NO: 8-13, 21-26 and 34-39 is due solely to conservative amino acid
10 substitutions and wherein said isolated polypeptide comprises at least one immunogenic fragment. In a preferred embodiment, the present invention encompasses an isolated polypeptide comprising an amino acid sequence homologous to an amino acid sequence selected from the group consisting of SEQ ID NO: 8-13, 21-26 and 34-39. These represent kidney cell surface
15 antigens.

The present invention also relates to an antibody that reacts with a polypeptide as disclosed herein, preferably a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO: 8-13,
20 21-26 and 34-39. Such an antibody may be polyclonal, monoclonal, recombinant or synthetic in origin.

In one such embodiment, said antibody is associated, either covalently or non-covalently, with a cytotoxic agent, for example, an
25 apoptotic agent. Thus, the present invention relates to an immunoconjugate comprising an antibody of the invention and a cytotoxic agent.

30 The present invention also relates to a process for treating cancer comprising contacting a cancerous cell with an agent having activity against an expression product encoded by a gene sequence selected from the group consisting of SEQ ID NO: 1-7, 14-20 and 27-33. In one such embodiment, the

cancerous cell is contacted *in vivo*. In another such embodiment, said agent has affinity for said expression product. In a preferred embodiment, such agent is an antibody disclosed herein, such as an antibody that is specific or selective for, or otherwise reacts with, a polypeptide of the invention. In a preferred embodiment, the expression product is a polypeptide incorporating an amino acid sequence selected from SEQ ID NO: 8-13, 21-26 and 34-39.

The present invention further encompasses an immunogenic composition comprising a polypeptide disclosed herein, as well as compositions formed using antibodies specific for these polypeptides.

The present invention is also directed to uses of such compositions. Such uses include a method for treating cancer in an animal afflicted therewith comprising administering to said animal an amount of an immunogenic composition of one or more of the polypeptides disclosed herein where such amount is an amount sufficient to elicit the production of cytotoxic T lymphocytes specific for a polypeptide of the invention, preferably a polypeptide incorporating a sequence of SEQ ID NO: 8-13, 21-26 and 34-39. In a preferred embodiment, the animal to be so treated is a human patient.

DEFINITIONS

As used herein, the terms "portion," "segment," and "fragment," when used in relation to polypeptides, refer to a continuous sequence of residues, such as amino acid residues, which sequence forms a subset of a larger sequence. For example, if a polypeptide were subjected to treatment with any of the common endopeptidases, such as trypsin or chymotrypsin, the oligopeptides resulting from such treatment would represent portions, segments or fragments of the starting polypeptide. When used in relation to a polynucleotides, such

terms refer to the products produced by treatment of said polynucleotides with any of the common endonucleases.

As used herein, the term "isolated" means that the material is removed from its original environment (e.g., the natural environment if it is naturally occurring). It could also be produced recombinantly and subsequently purified. For example, a naturally-occurring polynucleotide or polypeptide present in a living animal is not isolated, but the same polynucleotide or polypeptide, separated from some or all of the coexisting materials in the natural system, is isolated. Such polynucleotides, for example, those prepared recombinantly, could be part of a vector and/or such polynucleotides or polypeptides could be part of a composition, and still be isolated in that such vector or composition is not part of its natural environment. In one embodiment of the present invention, such isolated, or purified, polypeptide is useful in generating antibodies for practicing the invention, or where said antibody is attached to a cytotoxic or cytolytic agent, such as an apoptotic agent.

The term "percent identity" or "percent identical," when referring to a sequence, means that a sequence is compared to a claimed or described sequence after alignment of the sequence to be compared (the "Compared Sequence") with the described or claimed sequence (the "Reference Sequence"). The Percent Identity is then determined according to the following formula:

$$\text{Percent Identity} = 100 [1 - (C/R)]$$

wherein C is the number of differences between the Reference Sequence and the Compared Sequence over the length of alignment between the Reference Sequence and the Compared Sequence wherein (i) each base or amino acid in the Reference Sequence that does not have a corresponding aligned base or amino acid in the Compared Sequence and (ii) each gap in the Reference Sequence and (iii) each aligned base or amino acid in the Reference Sequence

that is different from an aligned base or amino acid in the Compared Sequence, constitutes a difference; and R is the number of bases or amino acids in the Reference Sequence over the length of the alignment with the Compared Sequence with any gap created in the Reference Sequence also being counted
5 as a base or amino acid.

If an alignment exists between the Compared Sequence and the Reference Sequence for which the percent identity as calculated above is about equal to or greater than a specified minimum Percent Identity then the
10 Compared Sequence has the specified minimum percent identity to the Reference Sequence even though alignments may exist in which the hereinabove calculated Percent Identity is less than the specified Percent Identity.

15 As known in the art "similarity" between two polypeptides is determined by comparing the amino acid sequence and its conserved amino acid substitutes of one polypeptide to the sequence of a second polypeptide.

In accordance with the present invention, the term "DNA segment" or
20 "DNA sequence" refers to a DNA polymer, in the form of a separate fragment or as a component of a larger DNA construct, which has been derived from DNA isolated at least once in substantially pure form, i.e., free of contaminating endogenous materials and in a quantity or concentration enabling identification, manipulation, and recovery of the segment and its
25 component nucleotide sequences by standard biochemical methods, for example, using a cloning vector. Such segments are provided in the form of an open reading frame uninterrupted by internal nontranslated sequences, or introns, which are typically present in eukaryotic genes. Sequences of non-translated DNA may be present downstream from the open reading frame,
30 where the same do not interfere with manipulation or expression of the coding regions.

The term "coding region" refers to that portion of a gene which either naturally or normally codes for the expression product of that gene in its natural genomic environment, i.e., the region coding *in vivo* for the native expression product of the gene. The coding region can be from a normal, mutated or altered gene, or can even be from a DNA sequence, or gene, wholly synthesized in the laboratory using methods well known to those of skill in the art of DNA synthesis.

In accordance with the present invention, the term "nucleotide sequence" refers to a heteropolymer of deoxyribonucleotides. Generally, DNA segments encoding the proteins provided by this invention are assembled from cDNA fragments and short oligonucleotide linkers, or from a series of oligonucleotides, to provide a synthetic gene which is capable of being expressed in a recombinant transcriptional unit comprising regulatory elements derived from a microbial, eukaryotic or viral operon.

The term "expression product" means that polypeptide or protein that is the natural translation product of the gene and any nucleic acid sequence coding equivalents resulting from genetic code degeneracy and thus coding for the same amino acid(s).

The term "active fragment," when referring to a coding sequence, means a portion comprising less than the complete coding region whose expression product retains essentially the same biological function or activity as the expression product of the complete coding region.

The term "primer" means a short nucleic acid sequence that is paired with one strand of DNA and provides a free 3'-OH end at which a DNA polymerase starts synthesis of a deoxyribonucleotide chain.

The term "promoter" means a region of DNA involved in binding of RNA polymerase to initiate transcription. The term "enhancer" refers to a region of DNA that, when present and active, has the effect of increasing expression of

a different DNA sequence that is being expressed, thereby increasing the amount of expression product formed from said different DNA sequence.

The term "open reading frame (ORF)" means a series of triplets coding
5 for amino acids without any termination codons and is a sequence (potentially) translatable into protein.

As used herein, reference to a "DNA sequence" includes both single
10 stranded and double stranded DNA. Thus, the specific sequence, unless the context indicates otherwise, refers to the single strand DNA of such sequence, the duplex of such sequence with its complement (double stranded DNA) and the complement of such sequence.

As used herein, "corresponding genes" refers to genes that encode an
15 RNA that is at least 90% identical, preferably at least 95% identical, most preferably at least 98% identical, and especially identical, to an RNA encoded by one of the nucleotide sequences disclosed herein (i.e., SEQ ID NO: 1-7, 14-20 and 27-33). Such genes will also encode the same polypeptide sequence as any of the sequences disclosed herein, preferably SEQ ID NO:
20 1-7, 14-20 and 27-33, but may include differences in such amino acid sequences where such differences are limited to conservative amino acid substitutions, such as where the same overall three dimensional structure, and thus the same antigenic character, is maintained. Thus, amino acid sequences may be within the scope of the present invention where they react
25 with the same antibodies that react with polypeptides comprising the sequences of SEQ ID NO: 8-13, 21-26 and 34-39. A "corresponding gene" includes splice variants thereof.

The genes identified by the present disclosure are considered "cancer-
30 related" genes, as this term is used herein, and include genes expressed at higher levels (due, for example, to elevated rates of expression, elevated extent of expression or increased copy number) in cancer cells relative to

expression of these genes in normal (i.e., non-cancerous) cells where said cancerous state or status of test cells or tissues has been determined by methods known in the art, such as by reverse transcriptase polymerase chain reaction (RT-PCR) as described in the Examples herein. In specific
5 embodiments, this relates to the genes whose sequences correspond to the sequences of SEQ ID NO: 1-7, 14-20 and 27-33.

As used herein, the term "conservative amino acid substitutions" are defined herein as exchanges within one of the following five groups:

- 10 I. Small aliphatic, nonpolar or slightly polar residues:
Ala, Ser, Thr, Pro, Gly;
- II. Polar, negatively charged residues and their amides:
Asp, Asn, Glu, Gln;
- 15 III. Polar, positively charged residues:
His, Arg, Lys;
- IV. Large, aliphatic, nonpolar residues:
Met Leu, Ile, Val, Cys
- V. Large, aromatic residues:
Phe, Tyr, Trp
- 20

DETAILED SUMMARY OF THE INVENTION

25 The present invention relates to processes for utilizing a nucleotide sequence for a cancer-linked gene, polypeptides encoded by such sequences and antibodies reactive with such polypeptides in methods of treating and diagnosing cancer, preferably kidney cancer, and in carrying out screening assays for agents effective in reducing the activity of cancer-linked genes and
30 thereby treating a cancerous condition.

The polypeptides disclosed herein incorporate various polynucleotide transcripts (SEQ ID NO: 1-7, 14-20 and 27-33) and the derived amino acid sequence (SEQ ID NO: 8-13, 21-26 and 34-39) from said transcripts are available as targets for chemotherapeutic agents, especially anti-cancer agents, including antibodies specific for said polypeptides.

The cancer-related polynucleotide sequences disclosed herein correspond to gene sequences whose expression is indicative of the cancerous status of a given cell. Such sequences are substantially identical to SEQ ID NO: 1-7, 14-20 and 27-33, which represent different transcripts identified from the GenBank EST database and which exhibit cancer-specific expression. The polynucleotides of the invention are those that correspond to a sequence of SEQ ID NO: 1-7, 14-20 and 27-33. Such sequences have been searched within the GenBank database, especially the EST database, with results as follows:

Type:	cell-surface tumor antigen therapeutic antibody target
Tissue:	kidney
Accession(s):	AI479935, AI479935, AI186520
Unigene cluster-ID(s):	Hs.61384
Chromosomal location:	3

The nucleotides and polypeptides, as gene products, used in the processes of the present invention may comprise a recombinant polynucleotide or polypeptide, a natural polynucleotide or polypeptide, or a synthetic polynucleotide or polypeptide, or a recombinant polynucleotide or polypeptide.

Fragments of such polynucleotides and polypeptides as are disclosed herein may also be useful in practicing the processes of the present invention.

For example, a fragment, derivative or analog of the polypeptide (SEQ ID NO: 8-13, 21-26 and 34-39) may be (i) one in which one or more of the amino acid residues are substituted with a conserved or non-conserved amino acid residue (preferably a conserved amino acid residue) and such substituted amino acid residue may or may not be one encoded by the genetic code, or (ii) one in which one or more of the amino acid residues includes a substituent group, or (iii) one in which the mature polypeptide is fused with another compound, such as a compound to increase the half-life of the polypeptide (for example, polyethylene glycol), or (iv) one in which the additional amino acids are fused to the mature polypeptide, such as a leader or secretory sequence or a sequence which is employed for purification of the mature polypeptide (such as a histidine hexapeptide) or a proprotein sequence. Such fragments, derivatives and analogs are deemed to be within the scope of those skilled in the art from the teachings herein.

15

In another aspect, the present invention relates to an isolated polypeptide, including a purified polypeptide, comprising an amino acid sequence at least 90% identical to the amino acid sequence of SEQ ID NO: 8-13 and/or 21-26 and/or 34-39. In preferred embodiments, said isolated polypeptide comprises an amino acid sequence having sequence identity of at least 95%, preferably at least about 98%, and especially is identical to, the sequence of SEQ ID NO: 8-13 and/or 21-26 and/or 34-39. The present invention also includes isolated active fragments of such polypeptides where said fragments retain the biological activity of the polypeptide or where such active fragments are useful as specific targets for cancer treatment, prevention or diagnosis. Thus, the present invention relates to any polypeptides, or fragments thereof, with sufficient sequence homology to the sequences disclosed herein as to be useful in the production of antibodies that react with (i.e., are selective or specific for) the polypeptides of SEQ ID NO: 8-13, 21-26 and 34-39 so as to be useful in targeting cells that exhibit such polypeptides, or fragments, on their surfaces, thereby providing targets for such antibodies and therapeutic agents associated with such antibodies.

The polynucleotides and polypeptides useful in practicing the processes of the present invention may likewise be obtained in an isolated or purified form. In addition, the polypeptide disclosed herein as being useful in practicing the processes of the invention are believed to be surface proteins present on cells, such as cancerous cells. Precisely how such cancer-linked proteins are used in the processes of the invention may thus differ depending on the therapeutic approach used. For example, cell-surface proteins, such as receptors, are desirable targets for cytotoxic antibodies that can be generated against the polypeptides disclosed herein.

The sequence information disclosed herein, as derived from the GenBank submissions, can readily be utilized by those skilled in the art to prepare the corresponding full-length polypeptide by peptide synthesis. The same is true for either the polynucleotides or polypeptides disclosed herein for use in the methods of the invention.

The present invention relates to an isolated polypeptide, encoded by one of the polynucleotide transcripts disclosed herein, comprising an amino acid sequence homologous to an amino acid sequence selected from the group consisting of SEQ ID NO: 8-13, 21-26 and 34-39, wherein any difference between amino acid sequence in the isolated polypeptide and the sequence of SEQ ID NO: 8-13, 21-26 and 34-39 is due solely to conservative amino acid substitutions and wherein said isolated polypeptide comprises at least one immunogenic fragment. In a preferred embodiment, the present invention encompasses an isolated polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO: 8-13, 21-26 and 34-39.

Methods of producing recombinant cells and vectors useful in preparing the polynucleotides and polypeptides disclosed herein are well known to those skilled in the molecular biology art. See, for example,

Sambrook, et al., *Molecular Cloning: A Laboratory Manual*, Second Edition, Cold Spring Harbor, N.Y., (1989), Wu et al., *Methods in Gene Biotechnology* (CRC Press, New York, NY, 1997), and *Recombinant Gene Expression Protocols*, in *Methods in Molecular Biology*, Vol. 62, (Tuan, ed., Humana Press, Totowa, NJ, 1997), the disclosures of which are hereby incorporated by reference.

In one aspect, the present invention relates to a process for identifying an agent that modulates the activity of a cancer-related gene comprising:

(a) contacting a compound with a cell containing a gene that corresponds to a polynucleotide having a sequence selected from the group consisting of SEQ ID NO: 1-7, 14-20 and 27-33 and under conditions promoting the expression of said gene; and

(b) detecting a difference in expression of said gene relative to when said compound is not present

thereby identifying an agent that modulates the activity of a cancer-related gene.

In specific embodiments of such process the cell is a cancer cell and the difference in expression is a decrease in expression. Such polynucleotides may also include those that have sequences identical to SEQ ID NO: 1-7, 14-20 and 27-33.

In another aspect, the present invention relates to a process for identifying an anti-neoplastic agent comprising contacting a cell exhibiting neoplastic activity with a compound first identified as a cancer related gene modulator using an assay process disclosed herein and detecting a decrease in said neoplastic activity after said contacting compared to when said contacting does not occur. Such neoplastic activity may include accelerated cellular replication and/or metastasis, and the decrease in neoplastic activity preferably results from the death of the cell.

The present invention also relates to a process for identifying an anti-neoplastic agent comprising administering to an animal exhibiting a cancer condition an effective amount of an agent first identified according to a process of one of one of the assays disclosed according to the invention and
5 detecting a decrease in said cancerous condition.

In specific embodiments of the present invention, the genes useful for the invention comprise genes that correspond to polynucleotides having a sequence selected from SEQ ID NO: 1-7, 14-20 and 27-33, or may comprise
10 the sequence of any of the polynucleotides disclosed herein (where the latter are cDNA sequences).

In accordance with the present invention, such assays rely on methods of determining the activity of the gene in question. Such assays are
15 advantageously based on model cellular systems using cancer cell lines, primary cancer cells, or cancerous tissue samples that are maintained in growth medium and treated with compounds at a single concentration or at a range of concentrations. At specific times after treatment, cellular RNAs are conveniently isolated from the treated cells or tissues, which RNAs are
20 indicative of expression of selected genes. The cellular RNA is then divided and subjected to differential analysis that detects the presence and/or quantity of specific RNA transcripts, which transcripts may then be amplified for detection purposes using standard methodologies, such as, for example, reverse transcriptase polymerase chain reaction (RT-PCR), etc. The presence
25 or absence, or concentration levels, of specific RNA transcripts are determined from these measurements. The polynucleotide sequences disclosed herein are readily used as probes for the detection of such RNA transcripts and thus the measurement of gene activity and expression.

30 The polynucleotides of the invention can include fully operational genes with attendant control or regulatory sequences or merely a polynucleotide

sequence encoding the corresponding polypeptide or an active fragment or analog thereof.

Because expression of the polynucleotide sequences disclosed herein
5 are specific to the cancerous state, useful gene modulation is downward modulation, so that, as a result of exposure to an antineoplastic agent identified by the screening assays herein, the corresponding gene of the cancerous cell is expressed at a lower level (or not expressed at all) when
10 exposed to the agent as compared to the expression when not exposed to the agent. For example, the gene sequences disclosed herein (SEQ ID NO: 1-7, 14-20 and 27-33) correspond to a gene expressed at a higher level in cells of kidney cancer than in normal kidney cells. Thus, where said chemical agent causes this gene of the tested cell to be expressed at a lower level than the same genes of the reference, this is indicative of downward modulation and
15 indicates that the chemical agent to be tested has anti-neoplastic activity.

In carrying out the assays disclosed herein, relative antineoplastic activity may be ascertained by the extent to which a given chemical agent modulates the expression of genes present in a cancerous cell. Thus, a first chemical agent
20 that modulates the expression of a gene associated with the cancerous state (i.e., a gene corresponding to one or more of the polynucleotide transcripts disclosed herein) to a larger degree than a second chemical agent tested by the assays of the invention is thereby deemed to have higher, or more desirable, or more advantageous, anti-neoplastic activity than said second chemical agent.

25

The gene expression to be measured is commonly assayed using RNA expression as an indicator. Thus, the greater the level of RNA (for example, messenger RNA or mRNA) detected the higher the level of expression of the corresponding gene. Thus, gene expression, either absolute or relative, is
30 determined by the relative expression of the RNAs encoded by such genes.

RNA may be isolated from samples in a variety of ways, including lysis and denaturation with a phenolic solution containing a chaotropic agent (e.g., trizol) followed by isopropanol precipitation, ethanol wash, and resuspension in aqueous solution; or lysis and denaturation followed by isolation on solid support, such as a Qiagen resin and reconstitution in aqueous solution; or lysis and denaturation in non-phenolic, aqueous solutions followed by enzymatic conversion of RNA to DNA template copies.

Normally, prior to applying the processes of the invention, steady state RNA expression levels for the genes, and sets of genes, disclosed herein will have been obtained. It is the steady state level of such expression that is affected by potential anti-neoplastic agents as determined herein. Such steady state levels of expression are easily determined by any methods that are sensitive, specific and accurate. Such methods include, but are in no way limited to, real time quantitative polymerase chain reaction (PCR), for example, using a Perkin-Elmer 7700 sequence detection system with gene specific primer probe combinations as designed using any of several commercially available software packages, such as Primer Express software., solid support based hybridization array technology using appropriate internal controls for quantitation, including filter, bead, or microchip based arrays, solid support based hybridization arrays using, for example, chemiluminescent, fluorescent, or electrochemical reaction based detection systems.

The gene expression indicative of a cancerous state need not be characteristic of every cell of a given tissue. Thus, the methods disclosed herein are useful for detecting the presence of a cancerous condition within a tissue where less than all cells exhibit the complete pattern. Thus, for example, a selected gene corresponding to the sequence of SEQ ID NO: 1, may be found, using appropriate probes, either DNA or RNA, to be present in as little as 60% of cells derived from a sample of tumorous, or malignant, tissue. In a highly preferred embodiment, such gene pattern is found to be

present in at least 100% of cells drawn from a cancerous tissue and absent from at least 100% of a corresponding normal, non-cancerous, tissue sample.

Expression of a gene may be related to copy number, and changes in
5 expression may be measured by determining copy number. Such change in
gene copy number may be determined by determining a change in expression
of messenger RNA encoded by a particular gene sequence, especially that of
SEQ ID NO: 1-7, 14-20 and 27-33. Also in accordance with the present
invention, said gene may be a cancer initiating or facilitating gene. In carrying
10 out the methods of the present invention, a cancer facilitating gene is a gene
that, while not directly initiating tumor formation or growth, acts, such as
through the actions of its expression product, to direct, enhance, or otherwise
facilitate the progress of the cancerous condition, including where such gene
acts against genes, or gene expression products, that would otherwise have
15 the effect of decreasing tumor formation and/or growth.

Although the expression of a gene corresponding to a sequence of
SEQ ID NO: 1-7, 14-20 and 27-33 may be indicative of a cancerous status for
a given cell, the mere presence of such a gene may not alone be sufficient to
20 achieve a malignant condition and thus the level of expression of such gene
may also be a significant factor in determining the attainment of a cancerous
state. Thus, it becomes essential to also determine the level of expression of
a gene as disclosed herein, including substantially similar sequences, as a
separate means of diagnosing the presence of a cancerous status for a given
25 cell, groups of cells, or tissues, either in culture or *in situ*.

The level of expression of the polypeptides disclosed herein is also a
measure of gene expression, such as polypeptides having sequence identical,
or similar to, any polypeptide encoded by a sequence of SEQ ID NO: 1-7, 14-
30 20 and 27-33, especially a polypeptide whose amino acid sequence is the
sequence of SEQ ID NO: 8-13, 21-26 and 34-39.

In accordance with the foregoing, the present invention specifically contemplates a method for determining the cancerous status of a cell to be tested, comprising determining the level of expression in said cell of a gene that includes one of the nucleotide sequences selected from the sequences of SEQ ID NO: 1-7, 14-20 and 27-33, including sequences substantially identical to said sequences, or characteristic fragments thereof, or the complements of any of the foregoing and then comparing said expression to that of a cell known to be non-cancerous whereby the difference in said expression indicates that said cell to be tested is cancerous.

10

In accordance with the invention, although gene expression for a gene that includes as a portion thereof one of the sequences of SEQ ID NO: 1-7, 14-20 and 27-33, is preferably determined by use of a probe that is a fragment of such nucleotide sequence, it is to be understood that the probe may be formed from a different portion of the gene. Expression of the gene may be determined by use of a nucleotide probe that hybridizes to messenger RNA (mRNA) transcribed from a portion of the gene other than the specific nucleotide sequence disclosed herein.

15

20

It should be noted that there are a variety of different contexts in which genes have been evaluated as being involved in the cancerous process. Thus, some genes may be oncogenes and encode proteins that are directly involved in the cancerous process and thereby promote the occurrence of cancer in an animal. In addition, other genes may serve to suppress the cancerous state in a given cell or cell type and thereby work against a cancerous condition forming in an animal. Other genes may simply be involved either directly or indirectly in the cancerous process or condition and may serve in an ancillary capacity with respect to the cancerous state. All such types of genes are deemed with those to be determined in accordance with the invention as disclosed herein. Thus, the gene determined by said process of the invention may be an oncogene, or the gene determined by said process may be a cancer facilitating gene, the latter including a gene that

25

30

directly or indirectly affects the cancerous process, either in the promotion of a cancerous condition or in facilitating the progress of cancerous growth or otherwise modulating the growth of cancer cells, either *in vivo* or *ex vivo*. In addition, the gene determined by said process may be a cancer suppressor gene, which gene works either directly or indirectly to suppress the initiation or progress of a cancerous condition. Such genes may work indirectly where their expression alters the activity of some other gene or gene expression product that is itself directly involved in initiating or facilitating the progress of a cancerous condition. For example, a gene that encodes a polypeptide, either wild or mutant in type, which polypeptide acts to suppress of tumor suppressor gene, or its expression product, will thereby act indirectly to promote tumor growth.

As noted previously, polynucleotides encoding the same proteins as any of SEQ ID NO: 1-7, 14-20 and 27-33, regardless of the percent identity of such sequences, are also specifically contemplated by any of the methods of the present invention that rely on any or all of said sequences, regardless of how they are otherwise described or limited. Thus, any such sequences are available for use in carrying out any of the methods disclosed according to the invention. Such sequences also include any open reading frames, as defined herein, present within the sequence of SEQ ID NO: 1-7, 14-20 and 27-33.

Because a gene disclosed according to the invention "corresponds to" a polynucleotide having a sequence of SEQ ID NO: 1-7, 14-20 and 27-33, said gene encodes an RNA (processed or unprocessed, including naturally occurring splice variants and alleles) that is at least 90% identical, preferably at least 95% identical, most preferably at least 98% identical to, and especially identical to, an RNA that would be encoded by, or be complementary to, such as by hybridization with, a polynucleotide having the indicated sequence. In addition, genes including sequences at least 90% identical to a sequence selected from SEQ ID NO: 1-7, 14-20 and 27-33, preferably at least about 95% identical to such a sequence, more preferably at

least about 98% identical to such sequence and most preferably comprising such sequence are specifically contemplated by all of the processes of the present invention. Sequences encoding the same proteins as any of these sequences, regardless of the percent identity of such sequences, are also specifically contemplated by any of the methods of the present invention that rely on any or all of said sequences, regardless of how they are otherwise described or limited. The polynucleotide sequences of the invention also include any open reading frames, as defined herein, present within any of the sequences of SEQ ID NO: 1-7, 14-20 and 27-33.

The sequences disclosed herein may be genomic in nature and thus represent the sequence of an actual gene, such as a human gene, or may be a cDNA sequence derived from a messenger RNA (mRNA) and thus represent contiguous exonic sequences derived from a corresponding genomic sequence, or they may be wholly synthetic in origin for purposes of practicing the processes of the invention. Because of the processing that may take place in transforming the initial RNA transcript into the final mRNA, the sequences disclosed herein may represent less than the full genomic sequence. They may also represent sequences derived from ribosomal and transfer RNAs. Consequently, the gene as present in the cell (and representing the genomic sequence) and the polynucleotide transcripts disclosed herein, including cDNA sequences, may be identical or may be such that the cDNAs contain less than the full genomic sequence. Such genes and cDNA sequences are still considered "corresponding sequences" (as defined elsewhere herein) because they both encode the same or related RNA sequences (i.e., related in the sense of being splice variants or RNAs at different stages of processing). Thus, by way of non-limiting example only, a gene that encodes an RNA transcript, which is then processed into a shorter mRNA, is deemed to encode both such RNAs and therefore encodes an RNA complementary to (using the usual Watson-Crick complementarity rules), or that would otherwise be encoded by, a cDNA (for example, a sequence as disclosed herein). Thus, the sequences disclosed herein correspond to genes

contained in the cancerous cells (here, kidney cancer) and are used to determine gene activity or expression because they represent the same sequence or are complementary to RNAs encoded by the gene. Such a gene also includes different alleles and splice variants that may occur in the cells

5 used in the methods of the invention, such as where recombinant cells are used to assay for anti-neoplastic agents and such cells have been engineered to express a polynucleotide as disclosed herein, including cells that have been engineered to express such polynucleotides at a higher level than is found in non-engineered cancerous cells or where such recombinant cells

10 express such polynucleotides only after having been engineered to do so. Such engineering includes genetic engineering, such as where one or more of the polynucleotides disclosed herein has been inserted into the genome of such cell or is present in a vector.

15 Such cells, especially mammalian cells, may also be engineered to express on their surfaces one or more of the polypeptides of the invention for testing with antibodies or other agents capable of masking such polypeptides and thereby removing the cancerous nature of the cell. Such engineering includes both genetic engineering, where the genetic complement of the cells

20 is engineered to express the polypeptide, as well as non-genetic engineering, whereby the cell has been physically manipulated to incorporate a polypeptide of the invention in its plasma membrane, such as by direct insertion using chemical and/or other agents to achieve this result.

25 In accordance with the foregoing, the present invention includes anti-cancer agents that are themselves either polypeptides, or small chemical entities, that affect the cancerous process, including initiation, suppression or facilitation of tumor growth, either *in vivo* or *ex vivo*. Said cancer modulating agent will have the effect of decreasing gene expression.

30

The present invention thus also relates to a method for treating cancer comprising contacting a cancerous cell with an agent having activity against

an expression product encoded by a gene or polynucleotide sequence as disclosed herein, such as one having, or corresponding to, the nucleotide sequence of SEQ ID NO: 1-7, 14-20 and 27-33. The present invention also relates to a process for treating cancer comprising contacting a cancerous cell with an agent having activity against an expression product encoded by a gene or polynucleotide sequence corresponding to a sequence selected from the group consisting of SEQ ID NO: 1-7, 14-20 and 27-33. In one such embodiment, the cancerous cell is contacted *in vivo*. In another such embodiment, said agent has affinity for said expression product. In a preferred embodiment, such agent is an antibody disclosed herein, such as an antibody that is specific or selective for, or otherwise reacts with, a polypeptide of the invention. In a preferred embodiment, the expression product is a polypeptide incorporating an amino acid sequence selected from SEQ ID NO: 8-13, 21-26 and 34-39.

15

The present invention is also directed to such uses of the compositions of polypeptides and antibodies disclosed herein. Such uses include a process for treating cancer in an animal afflicted therewith comprising administering to said animal an amount of an immunogenic composition of one or more of the polypeptides disclosed herein where such amount is an amount sufficient to elicit the production of cytotoxic T lymphocytes specific for a polypeptide of the invention, preferably a polypeptide incorporating a sequence of SEQ ID NO: 8-13, 21-26 and 34-39. In a preferred embodiment, the animal to be so treated is a human patient.

25

The proteins encoded by the genes disclosed herein due to their expression, or elevated expression, in cancer cells, represent highly useful therapeutic targets for "targeted therapies" utilizing such affinity structures as, for example, antibodies coupled to some cytotoxic agent. In such methodology, it is advantageous that nothing need be known about the endogenous ligands or binding partners for such cell surface molecules. Rather, an antibody or equivalent molecule that can specifically recognize the

30

cell surface molecule (which could include an artificial peptide, a surrogate ligand, and the like) that is coupled to some agent that can induce cell death or a block in cell cycling offers therapeutic promise against these proteins. Thus, such approaches include the use of so-called suicide "bullets" against intracellular proteins. For example, monoclonal antibodies may readily be produced by methods well known in the art, for example, the method of Kohler and Milstein (see: *Nature*, **256**:495 (1975)).

With the advent of methods of molecular biology and recombinant technology, it is now possible to produce antibody molecules by recombinant means and thereby generate gene sequences that code for specific amino acid sequences found in the polypeptide structure of the antibodies. Such antibodies can be produced by either cloning the gene sequences encoding the polypeptide chains of said antibodies or by direct synthesis of said polypeptide chains, with *in vitro* assembly of the synthesized chains to form active tetrameric (H_2L_2) structures with affinity for specific epitopes and antigenic determinants. This has permitted the ready production of antibodies having sequences characteristic of neutralizing antibodies from different species and sources.

Regardless of the source of the antibodies, or how they are recombinantly constructed, or how they are synthesized, *in vitro* or *in vivo*, using transgenic animals, such as cows, goats and sheep, using large cell cultures of laboratory or commercial size, in bioreactors or by direct chemical synthesis employing no living organisms at any stage of the process, all antibodies have a similar overall 3 dimensional structure. This structure is often given as H_2L_2 and refers to the fact that antibodies commonly comprise 2 light (L) amino acid chains and 2 heavy (H) amino acid chains. Both chains have regions capable of interacting with a structurally complementary antigenic target. The regions interacting with the target are referred to as "variable" or "V" regions and are characterized by differences in amino acid sequence from antibodies of different antigenic specificity.

The variable regions of either H or L chains contains the amino acid sequences capable of specifically binding to antigenic targets. Within these sequences are smaller sequences dubbed "hypervariable" because of their extreme variability between antibodies of differing specificity. Such
5 hypervariable regions are also referred to as "complementarity determining regions" or "CDR" regions. These CDR regions account for the basic specificity of the antibody for a particular antigenic determinant structure.

The CDRs represent non-contiguous stretches of amino acids within
10 the variable regions but, regardless of species, the positional locations of these critical amino acid sequences within the variable heavy and light chain regions have been found to have similar locations within the amino acid sequences of the variable chains. The variable heavy and light chains of all
15 antibodies each have 3 CDR regions, each non-contiguous with the others (termed L1, L2, L3, H1, H2, H3) for the respective light (L) and heavy (H) chains. The accepted CDR regions have been described by Kabat et al., *J. Biol. Chem.* **252**:6609-6616 (1977).

In all mammalian species, antibody polypeptides contain constant (i.e.,
20 highly conserved) and variable regions, and, within the latter, there are the CDRs and the so-called "framework regions" made up of amino acid sequences within the variable region of the heavy or light chain but outside the CDRs.

25 The antibodies disclosed according to the invention may also be wholly synthetic, wherein the polypeptide chains of the antibodies are synthesized and, possibly, optimized for binding to the polypeptides disclosed herein as being receptors. Such antibodies may be chimeric or humanized antibodies and may be fully tetrameric in structure, or may be dimeric and comprise only
30 a single heavy and a single light chain. Such antibodies may also include fragments, such as Fab and F(ab₂)' fragments, capable of reacting with and binding to any of the polypeptides disclosed herein as being receptors.

In one aspect, the present invention relates to immunoglobulins, or antibodies, as described herein, that react with, especially where they are specific for, the polypeptides having amino acid sequences as disclosed herein, preferably those having an amino acid sequence of one of SEQ ID NO: 8-13, 21-26 and 34-39. Such antibodies may commonly be in the form of a composition, especially a pharmaceutical composition. Such antibodies, by themselves, may have therapeutic value in that they are able to bind to, and thereby tie up, surface sites on cancerous cells. Where such sites have some type of function to perform (i.e., where they are surface enzymes, or channel structures, or structures that otherwise facilitate, actively or passively, the transport of nutrients and other vital materials to the cell. Such nutrients serve to facilitate the growth and replication of the cell and molecules that bind to such sites and thereby interfere with such activities can prove to have a therapeutic effect in that the result of such binding is to remove sources of nutrients from such cells, thereby interfering with growth and replication. In like manner, such binding may serve to remove vital enzyme activities from the cell's functional repertoire, thereby also interfering with viability and/or the ability of the cell to multiply or metastasize. In addition, by binding to such surface sites, the antibodies may serve to prevent the cells from reacting to environmental agents, such as cytokines and the like, that may facilitate growth, replication and metastasis, thereby further reducing the cancerous status of such cell and ameliorating the cancerous condition in a patient, even without proving fatal to the cell or cells so affected.

The methods of the present invention also include processes wherein the cancer cell is contacted *in vivo* as well as *ex vivo* with an agent that comprises a portion, or is part of an overall molecular structure, having affinity for an expression product of a gene corresponding to a polynucleotide sequence as disclosed herein, preferably where the expression product is a cell surface structure, most preferably a polypeptide as disclosed herein, such as one that comprises an amino acid sequence of SEQ ID NO: 8-13, 21-26 and 34-39. In one such embodiment, said portion having affinity for said

expression product is an antibody, especially where said expression product is a polypeptide or oligopeptide or comprises an oligopeptide portion, or comprises a polypeptide.

5 In another aspect, the present invention also relates to an antibody that reacts with a polypeptide as disclosed herein, preferably a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO: 8-13, 21-26 and 34-39. Such an antibody may be polyclonal, monoclonal, recombinant or synthetic in origin. In one such embodiment, said
10 antibody is associated, either covalently or non-covalently, with a cytotoxic agent, for example, an apoptotic agent. It is thus contemplated that the antibody acts a targeted vector for guiding an associated therapeutic agent to a cancerous cell, such as a cell expressing a polypeptide homologous to, if not identical to, a polypeptide as disclosed herein.

15

Where the cytotoxic agent is itself a polypeptide, said may be linked directly to an antibody specific for a surface target on a cancer cell, such as where the polypeptide represents an extension of the amino acid chain of the antibody. In alternative embodiments, such molecules may be covalently
20 linked through a linker sequence of long or short duration, such as an amino acid sequence of 5 to 10 residues in length. Where the cytotoxic agents is some small organic molecule, such as a small organic compound, or some type of apoptotic agent, this may be covalently bonded to the antibody molecule or may be attached by some other type of non-covalent linkage,
25 including hydrophobic and electrostatic linkages. Methods for forming such linkages, especially covalent linkages, are well known to those skilled in the art.

The antibodies disclosed herein may also serve as targeting vectors for
30 much larger structures, such as liposomes. In one such embodiment, an antibody is part of, or otherwise linked to, or associated with, a membranous structure, preferably a liposome or possibly some type of cellular organelle,

which acts as a reservoir for a cytotoxic agent, such as ricin. The antibody then acts to target said liposome to a cancerous tissue in an animal, whereupon the liposome provides a source of cytotoxic agents for localized treatment of a solid tumor or other type of neoplasm.

5

The present invention further encompasses an immunogenic composition comprising a polypeptide disclosed herein, as well as compositions formed using antibodies specific for these polypeptides.

10 Methods well known in the art for making formulations are found in, for example, *Remington: The Science and Practice of Pharmacy*, (19th ed.) Ed. A.R. Gennaro, 1995, Mack Publishing Company, Easton, PA. Formulations for parenteral administration may, for example, contain excipients, sterile water, or saline, polyalkylene glycols such as polyethylene glycol, oils of
15 vegetable origin, or hydrogenated naphthalenes. Biocompatible, biodegradable lactide polymer, lactide/glycolide copolymer, or polyoxyethylene-polyoxypropylene copolymers may be used to control the release of the compounds. Other potentially useful parenteral delivery systems for agonists of the invention include ethylenevinyl acetate copolymer particles, osmotic
20 pumps, implantable infusion systems, and liposomes. Formulations for inhalation may contain excipients, or example, lactose, or may be aqueous solutions containing, for example, polyoxyethylene-9-lauryl ether, glycocholate and deoxycholate, or may be oily solutions for administration in the form of nasal drops, or as a gel. It should be noted that, where the therapeutic agent
25 to be administered is an immunoconjugate, these sometimes contain chemical linkages that are somewhat labile in aqueous media and therefore must be stored prior to administration in a more stable environment, such as in the form of a lyophilized powder.

30 Such an agent can be a single molecular structure, comprising both affinity portion and anti-cancer activity portions, wherein said portions are derived from separate molecules, or molecular structures, possessing such

activity when separated and wherein such agent has been formed by combining said portions into one larger molecular structure, such as where said portions are combined into the form of an adduct. Said anti-cancer and affinity portions may be joined covalently, such as in the form of a single
5 polypeptide, or polypeptide-like, structure or may be joined non-covalently, such as by hydrophobic or electrostatic interactions, such structures having been formed by means well known in the chemical arts. Alternatively, the anti-cancer and affinity portions may be formed from separate domains of a single molecule that exhibits, as part of the same chemical structure, more than one
10 activity wherein one of the activities is against cancer cells, or tumor formation or growth, and the other activity is affinity for an expression product produced by expression of genes related to the cancerous process or condition.

In one embodiment of the present invention, a chemical agent, such as
15 a protein or other polypeptide, is joined to an agent, such as an antibody, having affinity for an expression product of a cancerous cell, such as a polypeptide or protein encoded by a gene related to the cancerous process, preferably a gene as disclosed herein according to the present invention, most preferably a polypeptide sequence disclosed herein. Thus, where the
20 presence of said expression product is essential to tumor initiation and/or growth, binding of said agent to said expression product will have the effect of negating said tumor promoting activity. In one such embodiment, said agent is an apoptosis-inducing agent that induces cell suicide, thereby killing the cancer cell and halting tumor growth.

25 Other genes within the cancer cell that are regulated in a manner similar to that of the genes disclosed herein and thus change their expression in a coordinated way in response to chemical compounds represent genes that are located within a common metabolic, signaling, physiological, or functional pathway so that by analyzing and identifying such commonly
30 regulated groups of genes (groups that include the gene, or similar sequences, disclosed according to the invention, one can (a) assign known genes and novel genes to specific pathways and (b) identify specific functions

and functional roles for novel genes that are grouped into pathways with genes for which their functions are already characterized or described. For example, one might identify a group of 10 genes, at least one of which is the gene as disclosed herein, that change expression in a coordinated fashion and for which the function of one, such as the polypeptide encoded by the sequence disclosed herein, is known then the other genes are thereby implicated in a similar function or pathway and may thus play a role in the cancer-initiating or cancer-facilitating process. In the same way, if a gene were found in normal cells but not in cancer cells, or happens to be expressed at a higher level in normal as opposed to cancer cells, then a similar conclusion may be drawn as to its involvement in cancer, or other diseases. Therefore, the processes disclosed according to the present invention at once provide a novel means of assigning function to genes, i.e. a novel method of functional genomics, and a means for identifying chemical compounds that have potential therapeutic effects on specific cellular pathways. Such chemical compounds may have therapeutic relevance to a variety of diseases outside of cancer as well, in cases where such diseases are known or are demonstrated to involve the specific cellular pathway that is affected.

The polypeptides disclosed herein, preferably those of SEQ ID NO: 8-13, 21-26 and 34-39, also find use as vaccines in that, where the polypeptide represents a surface protein present on a cancer cell, such polypeptide may be administered to an animal, especially a human being, for purposes of activating cytotoxic T lymphocytes (CTLs) that will be specific for, and act to lyse, cancer cells in said animal. Where used as vaccines, such polypeptides are present in the form of a pharmaceutical composition. The present invention may also employ polypeptides that have the same, or similar, immunogenic character as the polypeptides of SEQ ID NO: 8-13, 21-26 and 34-39 and thereby elicit the same, or similar, immunogenic response after administration to an animal, such as an animal at risk of developing cancer, or afflicted therewith. Thus, the polypeptides disclosed according to the invention will commonly find use as immunogenic compositions.

Expression of a gene corresponding to a polynucleotide disclosed herein, when in normal tissues, may indicate a predisposition towards development of kidney cancer. The encoded polypeptide might then present a potentially useful cell surface target for therapeutic molecules such as cytolytic antibodies, or antibodies attached to cytotoxic, or cytolytic, agents. .

The present invention specifically contemplates use of antibodies against the polypeptides encoded by the polynucleotides corresponding to the genes disclosed herein, whereby said antibodies are conjugates to one or more cytotoxic agents so that the antibodies serve to target the conjugated immunotoxins to a region of cancerous activity, such as a solid tumor. For many known cytotoxic agents, lack of selectivity has presented a drawback to their use as therapeutic agents in the treatment of malignancies. For example, the class of two-chain toxins, consisting of a binding subunit (or B-chain) linked to a toxic subunit (A-chain) are extremely cytotoxic. Thus, such agents as ricin, a protein isolated from castor beans, kills cells at very low concentrations (even less than 10^{-11} M) by inactivating ribosomes in said cells (see, for example, Lord et al., Ricin: structure, mode of action, and some current applications. *Faseb J*, 8: 201-208 (1994), and Blättler et al., Realizing the full potential of immunotoxins. *Cancer Cells*, 1: 50-55 (1989)). While isolated A-chains of protein toxins that functionally resemble ricin A-chain are only weakly cytotoxic for intact cells (in the concentration range of 10^{-7} to 10^{-6} M), they are very potent cytotoxic agents inside the cells. Thus, a single molecule of the A-subunit of diphtheria toxin can kill a cell once inside (see: Yamaizumi et al., One molecule of diphtheria toxin fragment A introduced into a cell can kill the cell. *Cell*, 15: 245-250, 1978).

The present invention solves this selectivity problem by using antibodies specific for antigens present on cancer cells to target the cytotoxins to said cells. In addition, use of antibodies decreases toxicity because the antibodies are non-toxic until they reach the tumor and, because the cytotoxin

is bound to the antibody, it is presented with less opportunity to cause damage to non-targeted tissues.

5 In addition, use of such antibodies alone can provide therapeutic effects on the tumor through the antibody-dependent cellular cytotoxic response (ADCC) and complement-mediated cell lysis mechanisms.

10 A number of recombinant immunotoxins (for example, consisting of Fv regions of cancer specific antibodies fused to truncated bacterial toxins) are well known (see, for example, Smyth et al., Specific targeting of chlorambucil to tumors with the use of monoclonal antibodies, *J. Natl. Cancer Inst.*, 76(3):503-510 (1986); Cho et al., Single-chain Fv/folate conjugates mediate efficient lysis of folate-receptor-positive tumor cells, *Bioconjug. Chem.*, 8(3):338-346 (1997)). As noted in the literature, these may contain, for
15 example, a truncated version of *Pseudomonas* exotoxin as a toxic moiety but the toxin is modified in such a manner that by itself it does not bind to normal human cells, but it retains all other functions of cytotoxicity. Here, recombinant antibody fragments target the modified toxin to cancer cells which are killed, such as by direct inhibition of protein synthesis, or by concomitant induction of
20 apoptosis. Cells that are not recognized by the antibody fragment, because they do not carry the cancer antigen, are not affected. Good activity and specificity has been observed for many recombinant immunotoxins in *in vitro* assays using cultured cancer cells as well as in animal tumor models. Ongoing clinical trials provide examples where the promising pre-clinical data
25 correlate with successful results in experimental cancer therapy. (see, for example, Brinkmann U., Recombinant antibody fragments and immunotoxin fusions for cancer therapy, *In Vivo* (2000) 14:21-27).

30 While the safety of employing immunoconjugates in humans has been established, *in vivo* therapeutic results have been less impressive. Because clinical use of mouse MAbs in humans is limited by the development of a foreign anti-globulin immune response by the human host, genetically

engineered chimeric human-mouse MAbs have been developed by replacing the mouse Fc region with the human constant region. In other cases, the mouse antibodies have been "humanized" by replacing the framework regions of variable domains of rodent antibodies by their human equivalents. Such humanized and engineered antibodies can even be structurally arranged to have specificities and effector functions determined by design and which characteristics do not appear in nature. The development of bispecific antibodies, having different binding ends so that more than one antigenic site can be bound, have proven useful in targeting cancer cells. Thus, such antibody specificity has been improved by chemical coupling to various agents such as bacterial or plant toxins, radionuclides or cytotoxic drugs and other agents. (see, for example, Bodey, B. et al). Genetically engineered monoclonal antibodies for direct anti-neoplastic treatment and cancer cell specific delivery of chemotherapeutic agents. *Curr Pharm Des* (2000) Feb;6(3):261-76). See also, Garnett, M. C., Targeted drug conjugates: principles and progress. *Adv. Drug Deliv. Rev.* (2001 Dec 17) 53(2):171-216; Brinkmann et al., Recombinant immunotoxins for cancer therapy. *Expert Opin Biol Ther.* (2001) 1(4):693-702.

Among the cytotoxic agents specifically contemplated for use as immunoconjugates according to the present invention are Calicheamicin, a highly toxic enediyne antibiotic isolated from *Micromonospora echinospora ssp. Calichensis*, and which binds to the minor groove of DNA to induce double strand breaks and cell death (see: Lee et al., Calicheamicins, a novel family of antitumor antibiotics. 1. Chemistry and partial structure of calicheamicin g₁. *J Am Chem Soc*, 109: 3464-3466 (1987); Zein et al., Calicheamicin gamma 1I: an antitumor antibiotic that cleaves double-stranded DNA site specifically, *Science*, 240: 1198-1201 (1988)). Useful derivatives of the calicheamicins include mylotarg and 138H11-Cam θ . Mylotarg is an immunoconjugate of a humanized anti-CD33 antibody (CD33 being found in leukemic cells of most patients with acute myeloid leukemia) and N-acetyl gamma colicheamicin dimethyl hydrazide, the latter of which is

readily coupled to an antibody of the present invention (in place of the anti-CD33 but which can also be humanized by substitution of human framework regions into the antibody during production as described elsewhere herein) to form an immunoconjugate of the invention. (see: Hamann et al. Gemtuzumab
5 Ozogamicin, A Potent and Selective Anti-CD33 Antibody-Calicheamicin Conjugate for Treatment of Acute Myeloid Leukemia, *Bioconjug. Chem.* **13**, 47-58 (2002)) For use with 138H11-Cam θ , 138H11 is an anti- γ -glutamyl transferase antibody coupled to theta calicheamicin through a disulfide linkage and found useful *in vitro* against cultured renal cell carcinoma cells.
10 (see: Knoll et al., Targeted therapy of experimental renal cell carcinoma with a novel conjugate of monoclonal antibody 138H11 and calicheamicin θ_1^1 , *Cancer Res*, **60**: 6089-6094 (2000) The same linkage may be utilized to link this cytotoxic agent to an antibody of the present invention, thereby forming a targeting structure for kidney cancer cells.

15

Also useful in forming the immunoconjugates of the invention is DC1, a disulfide-containing analog of adozelesin, that kills cells by binding to the minor groove of DNA, followed by alkylation of adenine bases. Adozelesin is a structural analog of CC-1065, an anti-tumor antibiotic isolated from microbial
20 fermentation of *Streptomyces zelensis*, and is about 1,000 fold more toxic to cultured cell lines than other DNA interacting agents, such as cis-platin and doxorubicin. This agent is readily linked to antibodies through the disulfide bond of adozelesin. (see: Chari et al., Enhancement of the selectivity and antitumor efficacy of a CC-1065 analogue through immunoconjugate
25 formation, *Cancer Res*, **55**: 4079-4084 (1995)).

Maytansine, a highly cytotoxic microtubular inhibitor isolated from the shrub *Maytenus serrata* found to have little value in human clinical trials, is much more effective in its derivatized form, denoted DM1, containing a disulfide bond to facilitate linkage to antibodies, is up to 10-fold more cytotoxic
30 (see: Chari et al., Immunoconjugates containing novel maytansinoids: promising anticancer drugs, *Cancer Res*, **52**: 127-131 (1992)). These same *in vitro* studies showed that up to four DM1 molecules could be linked to a single

immunoglobulin without destroying the binding affinity. Such conjugates have been used against breast cancer antigens, such as the *neu*/HER2/*erbB*-2 antigen. (see: Goldmacher et al., Immunogen, Inc., (2002) *in press*); also see Liu, C. et al., Eradication of large colon tumor xenografts by targeted delivery of maytansinoids, *Proc. Natl. Acad. Sci. USA*, **93**, 8618-8623 (1996)). For example, Liu et al. (1996) describes formation of an immunoconjugate of the maytansinoid cytotoxin DM1 and C242 antibody, a murine IgG1 immunoglobulin, available from Pharmacia and which has affinity for a mucin-like glycoprotein variably expressed by human colorectal cancers. The latter immunoconjugate was prepared according to Chari et al., *Cancer Res.*, **52**:127-131 (1992) and was found to be highly cytotoxic against cultured colon cancer cells as well as showing anti-tumor effects *in vivo* in mice bearing subcutaneous COLO 205 human colon tumor xenografts using doses well below the maximum tolerated dose.

15

In addition, there are a variety of protein toxins (cytotoxic proteins), which include a number of different classes, such as those that inhibit protein synthesis: ribosome-inactivating proteins of plant origin, such as ricin, abrin, gelonin, and a number of others, and bacterial toxins such as pseudomonas exotoxin and diphtheria toxin.

20

Another useful class is the one including taxol, taxotere, and taxoids. Specific examples include paclitaxel (taxol), its analog docetaxel (taxotere), and derivatives thereof. The first two are clinical drugs used in treating a number of tumors while the taxoids act to induce cell death by inhibiting the de-polymerization of tubulin. Such agents are readily linked to antibodies through disulfide bonds without disadvantageous effects on binding specificity.

25

In one instance, a truncated *Pseudomonas* exotoxin was fused to an anti-CD22 variable fragment and used successfully to treat patients with chemotherapy-resistant hairy-cell leukemia. (see: Kreitman et al., Efficacy of

30

the anti-CD22 recombinant immunotoxin BL22 in chemotherapy-resistant hairy-cell leukemia, *N Engl J Med*, **345**: 241-247 (2001)) Conversely, the cancer-linked peptides of the present invention offer the opportunity to prepare antibodies, recombinant or otherwise, against the appropriate
5 antigens to target solid tumors, preferably those of malignancies of kidney tissue, using the same or similar cytotoxic conjugates. Thus, many of the previously used immunoconjugates have been formed using antibodies against general antigenic sites linked to cancers whereas the antibodies formed using the peptides disclosed herein are more specific and target the
10 antibody-cytotoxic agent to a particular tissue or organ, thus further reducing toxicity and other undesirable side effects.

In addition, the immunoconjugates formed using the antibodies prepared against the cancer-linked antigens disclosed herein can be formed
15 by any type of chemical coupling. Thus, the cytotoxic agent of choice, along with the immunoglobulin, can be coupled by any type of chemical linkage, covalent or non-covalent, including electrostatic linkage, to form the immunoconjugates of the present invention.

When used as immunoconjugates, the antitumor agents of the present invention represent a class of pro-drugs that are relatively non-toxic when first administered to an animal (due mostly to the stability of the immunoconjugate), such as a human patient, but which are targeted by the conjugated immunoglobulin to a cancer cell where they then exhibit good
20 toxicity. The tumor-related, associated, or linked, antigens, preferably those presented herein, serve as targets for the antibodies (monoclonal, recombinant, and the like) specific for said antigens. The end result is the release of active cytotoxic agent inside the cell after binding of the immunoglobulin portion of the immunoconjugate.
25

30

The cited references describe a number of useful procedures for the chemical linkage of cytotoxic agents to immunoglobulins and the disclosures

of all such references cited herein are hereby incorporated by reference in their entirety. For other reviews see Ghetie et al., Immunotoxins in the therapy of cancer: from bench to clinic, *Pharmacol Ther*, **63**: 209-234 (1994), Pietersz et al. The use of monoclonal antibody immunoconjugates in cancer therapy, *Adv Exp Med Biol*, **353**:169-179 (1994), and Pietersz, G. A. The linkage of cytotoxic drugs to monoclonal antibodies for the treatment of cancer, *Bioconjug Chem*, **1**:89-95 (1990).

Thus, the present invention provides highly useful cancer-associated antigens for generation of antibodies for linkage to a number of different cytotoxic agents which are already known to have some *in vitro* toxicity and possess chemical groups available for linkage to antibodies.

The present invention also relates to a process that comprises a method for producing a product, such as test data, comprising identifying an agent according to one of the disclosed processes for identifying such an agent (i.e., the therapeutic agents identified according to the assay procedures disclosed herein) wherein said product is the data collected with respect to said agent as a result of said identification process, or assay, and wherein said data is sufficient to convey the chemical character and/or structure and/or properties of said agent. For example, the present invention specifically contemplates a situation whereby a user of an assay of the invention may use the assay to screen for compounds having the desired enzyme modulating activity and, having identified the compound, then conveys that information (i.e., information as to structure, dosage, etc) to another user who then utilizes the information to reproduce the agent and administer it for therapeutic or research purposes according to the invention. For example, the user of the assay (user 1) may screen a number of test compounds without knowing the structure or identity of the compounds (such as where a number of code numbers are used the first user is simply given samples labeled with said code numbers) and, after performing the screening process, using one or more assay processes of the present invention, then

imparts to a second user (user 2), verbally or in writing or some equivalent fashion, sufficient information to identify the compounds having a particular modulating activity (for example, the code number with the corresponding results). This transmission of information from user 1 to user 2 is specifically contemplated by the present invention.

It should be cautioned that, in carrying out the procedures of the present invention as disclosed herein, whether to form immunoconjugates or screen for other antitumor agents using the genes and polypeptides disclosed herein, any reference to particular buffers, media, reagents, cells, culture conditions and the like are not intended to be limiting, but are to be read so as to include all related materials that one of ordinary skill in the art would recognize as being of interest or value in the particular context in which that discussion is presented. For example, it is often possible to substitute one buffer system or culture medium for another and still achieve similar, if not identical, results. Those of skill in the art will have sufficient knowledge of such systems and methodologies so as to be able, without undue experimentation, to make such substitutions as will optimally serve their purposes in using the methods and procedures disclosed herein.

The present invention will now be further described by way of the following non-limiting example. In applying the disclosure of the example, it should be kept clearly in mind that other and different embodiments of the methods disclosed according to the present invention will no doubt suggest themselves to those of skill in the relevant art. The following example shows how a potential anti-neoplastic agent may be identified using one or more of the genes disclosed herein.

EXAMPLE

Determination of Gene Inhibitory Activity of an Anti-neoplastic Agent

5 SW480 cells are grown to a density of 10^5 cells/cm² in Leibovitz's L-15 medium supplemented with 2 mM L-glutamine (90%) and 10% fetal bovine serum. The cells are collected after treatment with 0.25% trypsin, 0.02% EDTA at 37°C for 2 to 5 minutes. The trypsinized cells are then diluted with 30 ml growth medium and plated at a density of 50,000 cells per well in a 96 well
10 plate (100 µl/well). The following day, cells are treated with either compound buffer alone, or compound buffer containing a chemical agent to be tested, for 24 hours. The media is then removed, the cells lysed and the RNA recovered using the RNeasy reagents and protocol obtained from Qiagen. RNA is quantitated and 10 ng of sample in 1 µl are added to 24 µl of Taqman reaction
15 mix containing 1X PCR buffer, RNAsin, reverse transcriptase, nucleoside triphosphates, amplitaq gold, tween 20, glycerol, bovine serum albumin (BSA) and specific PCR primers and probes for a reference gene (18S RNA) and a test gene (Gene X). Reverse transcription is then carried out at 48°C for 30 minutes. The sample is then applied to a Perlin Elmer 7700 sequence
20 detector and heat denatured for 10 minutes at 95°C. Amplification is performed through 40 cycles using 15 seconds annealing at 60°C followed by a 60 second extension at 72°C and 30 second denaturation at 95°C. Data files are then captured and the data analyzed with the appropriate baseline windows and thresholds.

25

The quantitative difference between the target and reference gene is then calculated and a relative expression value determined for all of the samples used. In this way, the ability of a chemotherapeutic agent to effectively and selectively reduce the activity of a cancer-specific gene is
30 readily ascertained. The overall expression of the cancer-specific gene, as modulated by one chemical agent relative to another, is also determined.

Chemical agents having the most effect in reducing gene activity are thereby identified as the most anti-neoplastic.

5

References:

- 10 Walter A. Blättler and Ravi Chari: Drugs to enhance the therapeutic potency of anti-cancer antibodies: antibody-drug conjugates as tumor-activated prodrugs. In *Anticancer Agents - Frontiers in Cancer Chemotherapy* (Iwao Ojima, Gregory D. Vite, Karl-Heinz Altmann, Eds.), American Chemical Society, pp. 317-338 (2001).
- 15 Dan L. Longo, Patricia L. Duffey, John G. Gribben, Elaine S. Jaffe, Brendan D. Curti, Barry L. Gause, John E. Janik, Virginia M. Braman, Dixie Esseltine, Wyndham H. Wilson, Dwight Kaufman, Robert E. Wittes, Lee M. Nadler, and Walter J. Urba: Combination chemotherapy followed by an Immunotoxin (Anti-B4-blocked Ricin) in patients with indolent lymphoma: results of a Phase II study. *Cancer J.* 6, 146-150 (2000).
- 20 Walter A. Blättler and John M. Lambert: Preclinical immunotoxin development. In *Monoclonal Antibody-Based Therapy of Cancer* (M. Grossbard, Ed.), Marcel Dekker, Inc. NY, NY, pp. 1-22 (1998).
- Ravi V. J. Chari: Targeted delivery of chemotherapeutics: tumor-activated prodrug therapy. In *Advanced Drug Delivery Reviews*, Elsevier Science B.V., pp. 89-104 (1998).
- 25 David T. Scadden, David P. Schenkein, Zale Bernstein, Barry Luskey, John Doweiko, Anil Tulpule, and Alexandra M. Levine: Immunotoxin combined with chemotherapy for patients with AIDS-related Non-Hodgkin's Lymphoma. *Cancer* 83, 2580-2587 (1998).
- 30 Changnian Liu and Ravi VJ Chari: The development of antibody delivery systems to target cancer with highly potent maytansinoids. *Exp. Opi. Invest. Drugs* 6, 169-172 (1997).
- 35 A. C. Goulet, Viktor S. Goldmacher, John M. Lambert, C. Baron, Dennis C. Roy and E. Kouassi: Conjugation of blocked ricin to an anti-CD19 monoclonal antibody increases antibody-induced cell calcium mobilization and CD19 internalization. *Blood* 90, 2364-2375 (1997).
- Changnian Liu, John M. Lambert, Beverly A. Teicher, Walter A. Blättler, and Rosemary O'Connor: Cure of multidrug-resistant human B-cell lymphoma xenografts by combinations of anti-B4-blocked ricin and chemotherapeutic drugs. *Blood* 87, 3892-3898 (1996).

- Rajeeva Singh, Lana Kats, Walter A. Blättler, and John M. Lambert: Formation of N-Substituted 2-Iminothiolanes when amino groups in proteins and peptides are modified by 2-Iminothiolane. *Anal. Biochem.* 236, 114-125 (1996).
- 5 Changnian Liu, B. Mitra Tadayoni, Lizabeth A. Bourret, Kristin M. Mattocks, Susan M. Derr, Wayne C. Widdison, Nancy L. Kedersha, Pamela D. Ariniello, Victor S. Goldmacher, John M. Lambert, Walter A. Blättler, and Ravi V.J. Chari: Eradication of large colon tumor xenografts by targeted delivery of maytansinoids. *Proc. Natl. Acad. Sci. USA* 93, 8618-8623 (1996).
- 10 Denis C. Roy, Sophie Ouellet, Christiane Le Houiller, Pamela D. Ariniello, Claude Perreault and John M. Lambert: Elimination of neuroblastoma and small-cell lung cancer cells with an anti-neural cell adhesion molecule immunotoxin. *J. Natl. Cancer Inst.* 88, 1136-1145 (1996).
- 15 Walter A. Blättler, Ravi V.J. Chari and John M. Lambert: Immunoconjugates. In *Cancer Therapeutics: Experimental and Clinical Agents*. (B. Teicher, Ed.), Humana Press, Totowa, NJ, pp. 371-394 (1996).
- 20 Michael L. Grossbard, John M. Lambert, Victor S. Goldmacher, Arnold S. Freedman, Jeanne Kinsella, Danny P. Ducello, Susan N. Rabinowe, Laura Elisea, Felice Carol, James A. Taylor, Walter A. Blättler, Carol L. Epstein, and Lee M. Nadler: Anti-B4-blocked Ricin: A phase I trial of 7 day continuous infusion in patients with B-cell neoplasms. *J. Clin. Oncol.* 11, 726-737 (1993).
- 25 Michael L. Grossbard, John G. Gribben, Arnold S. Freedman, John M. Lambert, Jeanne Kinsella, Susan N. Rabinowe, Laura Eliseo, James A. Taylor, Walter A. Blättler, Carol L. Epstein, and Lee M. Nadler: Adjuvant immunotoxin therapy with anti-B4-blocked ricin following autologous bone marrow transplantation for patients with B-cell Non-Hodgkin's lymphoma. *Blood* 81, 2263-2271 (1993).
- 30 Sudhir A. Shah, Patricia M. Halloran, Cynthia A. Ferris, Beth A. Levine, Lizabeth A. Bourret, Victor S. Goldmacher, and Walter A. Blättler: Anti-B4-blocked Ricin immunotoxin shows therapeutic efficacy in four different SCID mouse tumor models. *Cancer Res.* 53, 1360-1367 (1993).
- 35 Ravi V.J. Chari, Bridget A. Martell, Jonathan L. Gross, Sherilyn B. Cook, Sudhir A. Shah, Walter A. Blättler, Sara J. McKenzie, and Victor S. Goldmacher: Immunoconjugates containing novel maytansinoids: promising anti-cancer drugs. *Cancer Res.* 52, 127-131 (1992).
- John M. Lambert, Peter D. Senter, Annie Yau-Young, Walter A. Blättler, and Victor S. Goldmacher: Purified immunotoxins that are reactive with human lymphoid cells. *J. Biol. Chem.* 250, 12035-12041 (1985).

WHAT IS CLAIMED IS:

1. A process for identifying an agent that modulates the activity of a cancer-related gene comprising:

5 (a) contacting a compound with a cell containing a gene that corresponds to a polynucleotide having a sequence selected from the group consisting of SEQ ID NO: 1-7, 14-20 and 27-33 and under conditions promoting the expression of said gene; and

(b) detecting a difference in expression of said gene relative to when
10 said compound is not present

thereby identifying an agent that modulates the activity of a cancer-related gene.

2 The process of claim 1 wherein said gene has a sequence selected
15 from the group consisting of SEQ ID NO: 1-7, 14-20 and 27-33.

3. The process of claim 2 wherein the cell is a cancer cell and the difference in expression is a decrease in expression.

20 4. The process of claim 3 wherein said cancer cell is a kidney cancer cell.

5. A process for identifying an anti-neoplastic agent comprising contacting a cell exhibiting neoplastic activity with a compound first identified
25 as a cancer related gene modulator using the process of one of claim 1 and detecting a decrease in said neoplastic activity after said contacting compared to when said contacting does not occur.

6. The process of claim 5 wherein said neoplastic activity is
30 accelerated cellular replication.

7. The process of claim 5 wherein said decrease in neoplastic activity results from the death of the cell.

5 8. A process for identifying an anti-neoplastic agent comprising administering to an animal exhibiting a cancer condition an effective amount of an agent first identified according to the process of one of claim 1 and detecting a decrease in said cancerous condition.

10 9. A process for determining the cancerous status of a cell, comprising determining an increase in the level of expression in said cell of a gene that corresponds to a polynucleotide having a sequence selected from the group consisting of SEQ ID NO: 1-7, 14-20 and 27-33 wherein an elevated expression relative to a known non-cancerous cell indicates a cancerous state or potentially cancerous state.

15

10. The process of claim 9 wherein said elevated expression is due to an increased copy number.

20 11. An isolated polypeptide comprising an amino acid sequence homologous to an amino acid sequence selected from the group consisting of SEQ ID NO: 8-13, 21-26 and 34-39 wherein any difference between said amino acid sequence and the sequence of SEQ ID NO: 8-13, 21-26 and 34-39 is due solely to conservative amino acid substitutions and wherein said isolated polypeptide comprises at least one immunogenic fragment.

25

12. An isolated polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO: 8-13, 21-26 and 34-39.

30

13. An antibody that reacts with a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO: 8-13, 21-26 and 34-39.

5 14. The antibody of claim 13 wherein said antibody is a recombinant antibody.

15 15. The antibody of claim 13 wherein said antibody is a synthetic antibody.

10

16. The antibody of claim 13 wherein said antibody is a humanized antibody.

15 17. An immunoconjugate comprising the antibody of claim 13 and a cytotoxic agent.

18. The antibody of claim 17 wherein said cytotoxic agent is a member selected from the group consisting of a calicheamicin, a maytansinoid, an adozelesin, a cytotoxic protein, a taxol, a taxotere, a taxoid and DC1.

20

19. The immunoconjugate of claim 18 wherein said calicheamicin is calicheamicin γ_1^I , N-acetyl gamma calicheamicin dimethyl hydrazide or calicheamicin θ_1^I .

25 20. The immunoconjugate of claim 18 wherein said maytansinoid is DM1.

21. The immunoconjugate of claim 18 wherein said cytotoxic protein is ricin, abrin, gelonin, pseudomonas exotoxin or diphtheria toxin.

30

22. The immunoconjugate of claim 18 wherein said taxol is paclitaxel.

23. The immunoconjugate of claim 18 wherein said taxotere is docetaxel.

5 24. A process for treating cancer comprising contacting a cancerous cell *in vivo* with an agent having activity against an expression product encoded by a gene sequence selected from the group consisting of SEQ ID NO: 1-7, 14-20 and 27-33.

10 25. The process of claim 24 wherein said agent is the antibody of claim 13.

26. The process of claim 24 wherein said agent is an immunoconjugate of claim 17.

15 27. An immunogenic composition comprising a polypeptide of claim 11.

28. An immunogenic composition comprising a polypeptide of claim 12.

20 29. The process of claim 24 wherein said cancer is kidney cancer.

30. A process for treating cancer in an animal afflicted therewith comprising administering to said animal an amount of an immunogenic composition of claim 27 sufficient to elicit the production of cytotoxic T lymphocytes specific for the polypeptide of claim 11.

25 31. A process for treating cancer in an animal afflicted therewith comprising administering to said animal an amount of an immunogenic composition of claim 28 sufficient to elicit the production of cytotoxic T lymphocytes specific for the polypeptide of claim 12.

30 32. A process for treating a cancerous condition in an animal afflicted therewith comprising administering to said animal a therapeutically effective

amount of an agent first identified as having anti-neoplastic activity using the process of claim 8.

33. A process for protecting an animal against cancer comprising administering to an animal at risk of developing cancer a therapeutically effective amount of an agent first identified as having anti-neoplastic activity using the process of claim 8.

34. The process of claim 30 wherein said animal is a human being.

35. The process of claim 30 wherein said cancer is kidney cancer.

36. A method for producing test data with respect to the gene modulating activity of a compound comprising:

(a) contacting a compound with one or more cells containing a polynucleotide comprising a nucleotide sequence corresponding to a gene whose expression is increased in a cancerous cell over that in a non-cancerous cell or a gene whose expression is elevated in a non-cancerous cell over that in a cancerous cell under conditions wherein said polynucleotide is being expressed, and

(b) determining a change in expression of more than one of said polynucleotides, and

(c) producing test data with respect to the gene modulating activity of said compound based on an increase in the expression of the determined genes whose expression is otherwise elevated in a non-cancerous cell over that in a cancerous cell and a decrease in the expression of the determined genes whose expression is otherwise increased in a cancerous cell over that in a non-cancerous cell indicating gene modulating activity.

SEQUENCE LISTING

DT05 Rec'd PCT/PTO 30 NOV 2004

<110> Avalon Pharmaceuticals

<120> Cancer-Linked Gene as Target for Chemotherapy

<130> 689290-165

<140>

<141>

<150> US/60/385,505

<151> 2002-06-04

<160> 39

<170> PatentIn version 3.0

<210> 1

<211> 4567

<212> DNA

<213> Homo sapiens

<400> 1

gcggccgccc	cattcccaga	ccggccgcca	gcccattctgg	ttagctcccg	ccgctccgcg	60
ccgcccggga	gtcgggagcc	gcgggggaacc	gggcacctgc	acccgcctct	gggagtgagt	120
ggttccagct	ggtgcctggc	ctgtgtctct	tggatgccct	gtggcttcag	tccgtctcct	180
gttgccacc	acctcgtccc	tgggcccgcct	gataccccag	cccaacagct	aaggtgtgga	240
tggacagtag	ggggctggct	tctctcactg	gtcaggggtc	ttctcccctg	tctgcctccc	300
ggagctagga	ctgcagaggg	gcctatcatg	gtgcttgacg	gccccctggc	tgtctcgctg	360
ttgctgccc	gcctcacact	gctggtgtcc	cacctctcca	gtccccagga	tgtctccagt	420
gagcccagca	gtgagcagca	gctgtgcgcc	cttagcaagc	acccccaccgt	ggcctttgaa	480
gacctgcagc	cgtgggtctc	taacttcacc	taccctggag	cccgggattt	ctcccagctg	540
gctttggacc	cctccgggaa	ccagctcatc	gtgggagcca	ggaactacct	cttcagactc	600
agccttgcca	atgtctctct	tcttcaggcc	acagagtggg	cctccagtga	ggacacgcgc	660
cgctcctgcc	aaagcaaagg	gaagactgag	gaggagtgtc	agaactacgt	gcgagtccctg	720
atcgctcgcc	gccggaaggt	gttcatgtgt	ggaaccaatg	ccttttcccc	catgtgcacc	780
agcagacagg	tggggaacct	cagccggact	actgagaaga	tcaatggtgt	ggcccgcctgc	840
ccctatgacc	cacgccacaa	ctccacagct	gtcatctcct	cccaggggga	gctctatgca	900
gccacggtca	tcgacttctc	aggctcgggac	cctgccatct	accgcagcct	gggcagtggg	960
ccaccgcttc	gcactgccc	atataactcc	aagtggctta	atgagccaaa	cttcgtggca	1020
gcctatgata	ttgggctgtt	tgcatacttc	ttcctgcggg	agaacgcagt	ggagcacgac	1080
tgtggacgca	ccgtgtactc	tcgcgtggcc	cgcgtgtgca	agaatgacgt	ggggggccga	1140
ttcctgctgg	aggacacatg	gaccacattc	atgaaggccc	ggctcaactg	ctcccgccccg	1200
ggcgaggtcc	ccttctacta	taacgagctg	cagagtgcct	tccacttgcc	agagcaggac	1260
ctcatctatg	gagttttcac	aaccaacgta	aacagcatcg	cggcttctgc	tgtctgcgcc	1320
ttcaacctca	gtgctatctc	ccaggctttc	aatggcccat	ttcgctacca	ggagaacccc	1380
agggtgcct	ggctccccat	agccaacccc	atccccaatt	tccagtgtgg	caccctgcct	1440
gagaccggtc	ccaacgagaa	cctgacggag	cgcagcctgc	aggacgcgca	gcgcctcttc	1500
ctgatgagcg	aggccgtgca	gccggtgaca	cccagaccct	gtgtcaccca	ggacagcgtg	1560
cgcttctcac	acctcgtggt	ggacctgggtg	caggctaaag	acacgctcta	ccatgtactc	1620
tacattggca	ccgagtcggg	caccatcctg	aaggcgtgt	ccacggcgag	ccgcagcctc	1680
cacggctgct	acctggagga	gctgcacgtg	ctgccccccg	ggcgccgcga	gccccctgcgc	1740
agcctgcgca	tctgcacag	cgcccgcgcg	ctcttcgtgg	ggctgagaga	cggcgtcctg	1800
cgggtcccac	tggagagggt	cgccgcctac	cgcagccagg	gggcatgcct	gggggcccgg	1860
gaccggtact	gtggctggga	cgggaagcag	caactgtgca	gcacactcga	ggacagctcc	1920
aacatgagcc	tctggaccca	gaacatcacc	gcctgtcctg	tgcggaatgt	gacacgggat	1980
gggggcttcg	gcccattggtc	accatggcaa	ccatgtgagc	acttgatgg	ggacaactca	2040
ggctcttgcc	tgtgtcgagc	tcgatcctgt	gattccccctc	gaccccgctg	tgggggcctt	2100
gactgcctgg	ggccagccat	ccacatcgcc	aactgctcca	ggaatggggc	gtggacccccg	2160

tggtcatcgt	gggcgctgtg	cagcacgtcc	tgtggcatcg	gcttccaggt	ccgccagcga	2220
agttgcagca	accctgctcc	cegccacggg	ggccgcacat	gcgtgggcaa	gagccgggag	2280
gaacggttct	gtaatgagaa	cacgccttgc	ccgggtgcca	tcttctgggc	ttcctggggc	2340
tcttgagca	agtgcagcag	caactgtgga	gggggcatgc	agtcgcggcg	tcgggcctgc	2400
gagaacggca	actcctgcct	gggctgcggc	gtggagttca	agacgtgcaa	ccccgagggc	2460
tgccccgaag	tgcggcgcaa	caccccttgc	acgcctgtgc	tggccgtgaa	cgtgacgcag	2520
ggcggggcac	ggcaggagca	gcgggttcgc	ttcacctgcc	gcgcgcccct	tgcagaccgc	2580
cacggcctgc	agttcggcag	gagaaggacc	gagacgagga	cctgtcccgc	ggacggctcc	2640
ggctcctgcg	acaccgacgc	cctggtggag	gtcctcctgc	gcagcgggag	cacctccccg	2700
cacacggtga	gcgggggctg	ggccgccttg	ggcccgtggt	cgctcctgctc	ccgggactgc	2760
gagctgggct	tccgcgtccg	caagagaacg	tgcactaacc	cggagccccg	caacgggggc	2820
ctgccttgcg	tgggcgatgc	tgccgagtag	caggactgca	acccccaggc	ttgccagatt	2880
cggggtgctt	ggctcctgctg	gacctcatgg	tctccatgct	cagcttctctg	tgggtgggggt	2940
cactatcaac	gcacccgttc	ctgcaccagc	cccgcacctc	ccccaggtga	ggacatctgt	3000
ctcgggctgc	acacggagga	ggcactatgt	gccacacagg	cctgcccaga	aggctggtcg	3060
ccctggtctg	agtggagtaa	gtgcactgac	gacggagccc	agagccgaag	ccggcactgt	3120
gaggagctcc	tcccagggtc	cagcgcmtgt	gctggaacaa	gcagccagag	ccgcccctgc	3180
ccctacagcg	agattcscgt	catcctgcca	gcctccagca	tggaggaggc	caccgactgt	3240
gcagggttca	atctcatcca	cttgggtggc	acgggcatct	cctgcttctt	gggctctggg	3300
ctcctgaccc	tagcagtgtg	cctgtcttgc	cagcactgcc	agcgtcagtc	ccaggagtcc	3360
acactggtcc	atcctgccac	ccccaacctc	ttgcactaca	agggcggagg	caccccgaag	3420
aatgaaaagt	acacacccat	ggaattcaag	accctgaaca	agaataactt	gatccctgat	3480
gacagagcca	acttctatcc	attgcagcag	accaatgtgt	acacgactac	ttactaccca	3540
agccccctga	acaaacacag	cttcgggccc	gaggcctcac	ctggacaacg	gtgcttcccc	3600
aacagctgat	accgccgtcc	tggggacttg	ggcttcttgc	cttcataagg	cacagagcag	3660
atggagatgg	gacagtggag	ccagtttggg	tttctccctc	tgcactaggc	caagaacttg	3720
ctgccttgcg	tgtggggggg	cccattccgg	ttcagagagc	tctggctggc	attgaccatg	3780
ggggaaaggg	ctgggtttcag	gctgacatat	ggccgcaggt	ccagttcagc	ccaggctctmt	3840
catggttatc	ttccaaccca	ctgtcacgct	gacactatgc	tgccatgcct	gggctgtgga	3900
cctactgggc	atgttgaggaa	ytggagaatg	gagatggcaa	gagggcaggc	ttttaagtgt	3960
gggttgagga	caacttcctg	tggccccac	aagctgagtc	tggccttctc	cagctggccc	4020
caaaaaaggc	ctttgctaca	tctgattat	ctctgaaagt	aatcaatcaa	gtggctccag	4080
tagctctgga	ttttctgcca	gggctgggccc	attgtggtgc	tggccccagta	tgacatggga	4140
ccaaggccag	cgcagggttat	ccacctctgc	ctggaagtct	atactctacc	cagggcatcc	4200
ctctggtcag	aggcagttag	tactgggaac	tggaggctga	cctgtgctta	gaagtccttt	4260
aatctgggct	ggtacaggcc	tcagccttgc	cctcaatgca	cgaaagggtg	cccaggagag	4320
aggatcaatg	ccataggagg	cagaagtctg	gcctctgtgc	ctctatggag	actatcttcc	4380
agttgctgct	caacagagtt	gttggctgag	acctgcttgg	gagtctctgc	tggcccttca	4440
tctgttcagg	aacacacaca	cacacacact	cacacacgca	cacacaatca	caatttgcta	4500
cagcaacaaa	aaagacattg	ggctgtggca	ttattaatta	aagatgatat	ccagtcaaaa	4560
aaaaact						4567

<210> 2
 <211> 453
 <212> DNA
 <213> Homo sapiens

<400> 2						
agtaatcagc	tcgggtaccgg	catgtgctgt	agccagcgca	ggttatccac	ctctgcctgg	60
aagtctatac	tctacccagg	gcateccctc	ggtcagaggc	agtgagtact	gggaactgga	120
ggctgacctg	tgccttagaag	tcctttaatc	tgggctggta	caggcctcag	ccttgccctc	180
aatgcacgaa	aggtggccca	ggagagagga	tcaatgccat	aggaggcaga	agtctggcct	240
ctgtgcctct	atggagacta	tcttccagtt	gctgctcaac	agagttgttg	gctgagacct	300
gcttgggagt	ctctgctggc	ccttcatctg	ttcaggaaca	cacacacaca	cacactcaca	360
cacgcacaca	caatcacaat	ttgctacagc	aacaaaaaag	acattgggct	gtggcattat	420
taattaaaga	tgatatccag	tcaaaaaaaa	act			453

<210> 3
 <211> 4675

<212> DNA

<213> Homo sapiens

<400> 3

gcggccgccc	cattcccaga	ccggccgcca	gcccattctgg	ttagctcccg	ccgctccgcg	60
ccgcccggga	gtcgggagcc	gcgggggaacc	gggcacctgc	accgcctct	gggagtgagt	120
ggttccagct	ggtgcctggc	ctgtgtctct	tggatgccct	gtggttcag	tccgtctcct	180
gttgcccacc	acctcgtccc	tgggcccgcct	gataccccag	cccaacagct	aagggtgtgga	240
tggacagtag	ggggctggct	tctctcactg	gtcaggggca	ggaagccaag	tggagacttt	300
gaatggaggc	aaggatggag	gggacctggg	gaagaggact	ggcctgaatc	accttcccca	360
aaggctctca	tggactgagc	tggaggctctt	ctccccctgtc	tgcctcccgg	agctaggact	420
gcagaggggc	ctatcatggt	gcttgccaggc	cccctggctg	tctcgtgtgt	gctgcccagc	480
ctcacactgc	tgggtgtccca	cctctccagc	tcccaggatg	tctccagtga	gcccagcagt	540
gagcagcagc	tgtgcgcct	tagcaagcac	cccaccgtgg	cctttgaaga	cctgcagccg	600
tgggtctcta	acttcaccta	ccctggagcc	ccggatttct	cccagctggc	tttggacccc	660
tccgggaacc	agctcatcgt	gggagccagg	aactacctct	tcagactcag	ccttgccaat	720
gtctctcttc	ttcaggccac	agagtgggcc	tccagtggag	acacgcgcgc	ctcctgccaa	780
agcaaaggga	agactgagga	ggagtgtcag	aactacgtgc	gagtcctgat	cgtcgcgggc	840
cggaagggtgt	tcatgtgtgg	aaccaatgcc	ttttcccca	tgtgcaccag	cagacagggtg	900
gggaacctca	gccggactac	tgagaagatc	aatggtgtgg	cccgtgccc	ctatgaccca	960
cgccacaact	ccacagctgt	catctcctcc	cagggggagc	tctatgcagc	cacggctcatc	1020
gacttctcag	gtcgggaacc	tgccatctac	cgcagcctgg	gcagtgggcc	accgcttcgc	1080
actgccaat	ataactccaa	gtggcttaat	gagccaaact	tcgtggcagc	ctatgatatt	1140
gggtcttttg	catacttctt	cctgcgggag	aacgcagtgg	agcacgactg	tggacgcacc	1200
gtgtactctc	gcgtggcccc	cgtgtgcaag	aatgacgtgg	ggggccgatt	cctgctggag	1260
gacacatgga	ccacattcat	gaaggccccg	ctcaactgct	cccggccggg	cgagggtcccc	1320
ttctactata	acgagctgca	gagtgccttc	cacttgccag	agcaggacct	catctatgga	1380
gttttcacaa	ccaacgtaaa	cagcatcgcg	gcttctgctg	tctgcgcctt	caacctcagt	1440
gctatctccc	aggctttcaa	tggcccattt	cgctaccagg	agaaccccag	ggctgcctgg	1500
ctccccatag	ccaaccccat	ccccaatctc	cagtgtggca	ccctgcctga	gaccggtccc	1560
aacgagaacc	tgacggagcg	cagcctgcag	gacgcgcagc	gcctcttctc	gatgagcgag	1620
cccgtgcagc	cggtgacacc	cgagccctgt	gtcaccagcg	acagcgtgcg	ctcttcacac	1680
ctcgtggtgg	acctggtgca	ggctaaagac	acgctctacc	atgtactcta	cattggcacc	1740
gagtcgggca	ccatcctgaa	ggcgtgtcc	acggcgagcc	gcagcctcca	cggtgctac	1800
ctggaggagc	tgcacgtgct	gccccccggg	cgccgcgagc	ccctgcgcag	cctgcgcac	1860
ctgcacagcg	cccgcgcgct	cttcgtgggg	ctgagagacg	gcgtcctgcg	ggtcccactg	1920
gagaggtgcg	ccgcctaccg	cagccagggg	gcatgcctgg	gggcccggga	cccgtactgt	1980
ggctgggacg	ggaagcagca	acgttgacgc	acactcgagg	acagctccaa	catgagcctc	2040
tggaccaga	acatcacccg	ctgtcctgtg	cggaatgtga	cacgggatgg	gggcttcggc	2100
ccatggtcac	catggcaacc	atgtgagcac	ttggatgggg	acaactcagg	ctcttgcctg	2160
tgtcgagctc	gatacctgta	ttcccctcga	ccccgctgtg	ggggccttga	ctgcctgggg	2220
ccagccatcc	acatcgccaa	ctgctccagg	aatggggcgt	ggaccccgtg	gtcatcgtgg	2280
gcgctgtgca	gcacgtcctg	tggcatcggc	ttccaggctc	gccagcgaag	ttgcagcaac	2340
cctgctcccc	gccacggggg	ccgcatctgc	gtgggcaaga	gccgggagga	acggttctgt	2400
aatgagaaca	cgccttgccc	ggtgcccac	ttctgggctt	cctggggctc	ctggagcaag	2460
tgcagcagca	actgtggagg	gggcatgcag	tcgcggcgct	gggcctgcga	gaacggcaac	2520
tcctgcctgg	gctgcggcgt	ggagtccaag	acgtgcaacc	ccgagggctg	ccccgaagtg	2580
cggcgcaaca	ccccctggac	gccgtggctg	cccgtgaacg	tgacgcaggg	cggggcacgg	2640
caggagcagc	ggttccgctt	cacctgccgc	gcgccccctg	cagaccgcga	cggcctgcag	2700
ttcggcagga	gaaggaccga	gacgaggacc	tgtcccgcgg	acggctccgg	ctcctgcgac	2760
accgacgccc	tgggtggagg	cctcctgcgc	agcgggagca	cctccccgca	cacgggtgagc	2820
gggggctggg	ccgcctgggg	cccgtggtcg	tcctgctccc	gggactgcga	gctgggcttc	2880
cgcgtccgca	agagaacgtg	cactaaccgc	gagccccgca	acgggggcct	gccctgcgtg	2940
ggcgatgctg	ccgagtacca	ggactgcaac	ccccaggctt	gccagttcg	gggtgcttgg	3000
tcctgctgga	cctcatggtc	tccatgctca	gcttctctgt	gtgggggtca	ctatcaacgc	3060
accggttctt	gcaccagccc	cgcacctcc	ccagggtgag	acatctgtct	cgggctgcac	3120
acggaggagg	cactatgtgc	cacacaggcc	tgcccagaag	gctggtcgcc	ctggtctgag	3180
tggagtaagt	gcactgacga	cggagcccag	agccgaagcc	ggcactgtga	ggagctcctc	3240
ccagggtcca	gcgcmgtgtg	tggaaacagc	agccagagcc	gcccctgccc	ctacagcgag	3300
attcscgtca	tcctgccagc	ctccagcatg	gaggaggcca	ccgactgtgc	aggggtcaat	3360
ctcatccact	tgggtggccac	gggcatctcc	tgcttcttgg	gctctgggct	cctgacccta	3420

gcagtgtacc	tgtcttgcca	gcactgccag	cgtcagtccc	aggagtccac	actggtccat	3480
cctgccaccc	ccaaccattt	gcactacaag	ggcggaggca	ccccgaagaa	tgaaaagtac	3540
acacccatgg	aattcaagac	cctgaacaag	aataacttga	tccttgatga	cagagccaac	3600
ttctacccat	tgcagcagac	caatgtgtac	acgactactt	actacccaag	ccccctgaac	3660
aaacacagct	tccggcccga	ggcctcacct	ggacaacggg	gcttcccca	cagctgatac	3720
cgccgtcctg	gggacttggg	cttcttgcc	tcataaggca	cagagcagat	ggagatggga	3780
cagtggagcc	agtttggttt	tctccctctg	cactaggcca	agaacttgct	gccttgcc	3840
tggggggtcc	catccggctt	cagagagctc	tggctggcat	tgaccatggg	ggaaagggct	3900
ggtttcaggc	tgacatatgg	ccgcaggctc	agttcagccc	aggtctmtca	tggttatctt	3960
ccaacccact	gtcacgctga	cactatgctg	ccatgcctgg	gctgtggacc	tactgggcat	4020
ttgaggaayt	ggagaatgga	gatggcaaga	gggcaggctt	ttaagtttgg	gttgagaca	4080
acttccctgtg	gccccacaa	gctgagctctg	gccttctcca	gctggcccca	aaaaaggcct	4140
ttgctacatc	ctgattatct	ctgaaagtaa	tcaatcaagt	ggctccagta	gctctggatt	4200
ttctgccagg	gctgggccat	tgtgggtgctg	ccccagtagt	acatgggacc	aaggccagcg	4260
cagggtatcc	acctctgcct	ggaagtctat	actctacca	gggcatccct	ctggtcagag	4320
gcagtgtgta	ctgggaactg	gaggctgacc	tgtgcttaga	agtcctttaa	tctgggctgg	4380
tacaggcctc	agccttgccc	tcaatgcacg	aaaggtggcc	caggagagag	gatcaatgcc	4440
ataggaggca	gaagtctggc	ctctgtgcct	ctatggagac	tatcttccag	ttgctgtca	4500
acagagttgt	tggctgagac	ctgcttggga	gtctctgctg	gcccttcac	tggtcaggaa	4560
cacacacaca	cacacactca	cacacgcaca	cacaatcaca	atttgctaca	gcaacaaaa	4620
agacattggg	ctgtggcatt	attaattaaa	gatgatatcc	agtcaaaaa	aaact	4675

<210> 4
 <211> 4731
 <212> DNA
 <213> Homo sapiens

<400>	4					
attggagatg	ctcgggggca	ggctgcccgg	ttgtgtcctg	cttttctg	gccagaccaa	60
gccgtctgga	gctgctggc	aggttttctt	gctgacctca	cctgaccaca	gtggcctggg	120
tggactctac	agggaaatgt	tgttttctcc	ctgggagcag	tagcagcagt	cctggctccc	180
ctggactgag	aactcctcat	cagccccagg	aagcccggac	cccctttcag	ggatctggaa	240
ccgggtgtgcc	tgtggcccca	ggtctgctcc	caggcgtggg	ctgaagtctc	gacttctgtc	300
gctgggggca	aggagtggga	gagcccagct	gctgcctggg	ctttggcaga	cagcaggctg	360
atggtgctgg	cttccccgag	actgcttctc	ctgcctgctg	tctgatttcc	ctgcatgggtg	420
cccgagctg	agctgctacg	ggtcttctcc	cctgtctgcc	tcccggagct	aggactgcag	480
aggggcctat	catggtgctt	gcaggccccc	tggctgtctc	gctgttgctg	cccagcctca	540
cactgctggt	gtcccacctc	tccagctccc	aggatgtctc	cagttagccc	agcagtggc	600
agcagctgtg	cgcccttagc	aagcacccca	ccgtggcctt	tgaagacctg	cagccgtggg	660
tctctaactt	cacctaccct	ggagcccggg	atttctccca	gctggctttg	gacccctccg	720
ggaaccagct	catctgtgga	gccaggaact	acctcttcag	actcagcctt	gccaatgtct	780
ctcttcttca	ggccacagag	tgggcctcca	gtgaggacac	gcgcccgtcc	tgccaaagca	840
aagggaagac	tgaggaggag	tgtcagaact	acgtgcgagt	cctgatcgct	gccggccgga	900
aggtgttcat	gtgtggaacc	aatgcctttt	ccccatgtg	caccagcaga	caggtgggga	960
acctcagccg	gactactgag	aagatcaatg	gtgtggcccg	ctgcccctat	gacccacgcc	1020
acaactccac	agctgtcatc	tctcccagg	gggagctcta	tgagccacg	gtcatcgact	1080
tctcaggtcg	ggaccctgcc	atctaccgca	gcctgggcag	tgggccaccg	cttcgcactg	1140
cccaatataa	ctccaagtgg	cttaatgagc	caaacttcgt	ggcagcctat	gatattgggc	1200
tgtttgcata	cttcttctctg	cgggagaacg	cagtggagca	cgactgtgga	cgcaccgtgt	1260
actctcgctg	ggcccgctg	tgcaagaatg	acgtgggggg	ccgattcctg	ctggaggaca	1320
catggaccac	attcatgaag	gcccggctca	actgctccc	cccgggagag	gtccccttct	1380
actataacga	gctgcagagt	gccttccact	tgccagagca	ggacctcatc	tatggagttt	1440
tcacaaccaa	cgtaaacagc	atcgcggtt	ctgctgtctg	cgccttcaac	ctcagtgtca	1500
tctcccaggc	tttcaatggc	ccatttcgct	accaggagaa	ccccagggt	gcctggctcc	1560
ccatagccaa	ccccatcccc	aatttccagt	gtggcaccct	gcctgagacc	gggtcccaacg	1620
agaacctgac	ggagcgcagc	ctgcaggacg	cgcagcgctt	cttctctgatg	agcgaggccg	1680
tgcagccggt	gacacccgag	ccctgtgtca	cccagacag	cgtgcgcttc	tcacacctcg	1740
tgggtgacct	ggtgcaggct	aaagacacgc	tctaccatgt	actctacatt	ggcaccgagt	1800
cgggcaccat	ctggaaggcg	ctgtccacgc	cgagccgcag	cctccacggc	tgctacctgg	1860
aggagctgca	cgtgctgccc	cccgggcgc	gcgagccctt	gcgagcctg	cgcacctctg	1920

acagcgcccg	cgcgctcttc	gtggggctga	gagacggcgt	cctgcggggtc	ccactggaga	1980
ggtgcgcgcg	ctaccgcagc	cagggggcat	gcctgggggc	ccgggacccg	tactgtggct	2040
gggacgggaa	gcagcaacgt	tgcagcacac	tcgaggacag	ctccaacatg	agcctctgga	2100
cccagaacat	caccgcctgt	cctgtgcgga	atgtgacacg	ggatgggggc	ttcggcccat	2160
ggtcaccatg	gcaaccatgt	gagcaacttg	atggggacaa	ctcaggctct	tgcctgtgtc	2220
gagctcgatc	ctgtgattcc	cctcgacccc	gctgtggggg	ccttgactgc	ctggggccag	2280
ccatccacat	cgccaactgc	tccaggaatg	gggctgtggc	cccgtggtca	tcgtgggcgc	2340
tgtgcagcac	gtcctgtggc	atcggttcc	aggtccgcca	gcgaagtgc	agcaaccctg	2400
ctccccgcca	cgggggcccgc	atctgcgtgg	gcaagagccg	ggaggaacgg	ttctgtaatg	2460
agaacacgcc	ttgcccgggtg	cccatcttct	gggcttccctg	gggctcctgg	agcaagtgc	2520
gcagcaactg	tggagggggc	atgcagtcgc	ggcgtcgggc	ctgcgagaac	ggcaactcct	2580
gcttggggtg	cggcgtggag	ttcaagacgt	gcaacccgga	gggctgcccc	gaagtgcggc	2640
gcaacacccc	ctggaccccg	tggctgcccgc	tgaacgtgac	gcagggcggg	gcacggcagg	2700
agcagcgggt	ccgcttcacc	tgccgcgcgc	cccttgagaga	cccgcacggc	ctgcagttcg	2760
gcaggagaag	gaccgagacg	aggacctgtc	ccgcggacgg	ctccggctcc	tgcgacaccg	2820
acgccctggg	ggaggtcctc	ctgcgcagcg	ggagcacctc	cccgcacacg	gtgagcgggg	2880
gctgggcccgc	ctggggcccgc	tggctgtcct	gctcccggga	ctgcgagctg	ggcttccgcg	2940
tccgcaagag	aacgtgcact	aaccgggagc	cccgcacacg	gggctgccc	tgcgtgggcg	3000
atgctgccga	gtaccaggac	tgcaaccccc	aggcttgccc	agttcggggg	gcttgggtcct	3060
gctggacctc	atggtctcca	tgtctagctt	cctgtgggtg	gggtcactat	caacgcaccc	3120
gttctctgac	cagccccgca	ccctccccgc	gtgaggacat	ctgtctcggg	ctgcacacgg	3180
aggaggcact	atgtgccaca	caggcctggc	cagaaggctg	gtcgccctgg	tctgattgga	3240
gtaagtgcac	tgagcagcga	gcccagagcc	gaagccggca	ctgtgaggag	ctcctcccag	3300
ggtccagcgc	mtgtgctgga	aacagcagcc	agagccgccc	ctgcccctac	agcgagattc	3360
scgtcatcct	gccagcctcc	agcatggagg	aggccaccga	ctgtgcaggg	ttcaatctca	3420
tccacttggg	ggccacgggc	atctcctgct	tcttgggctc	tgggctcctg	accctagcag	3480
tgtacctgtc	ttgccagcac	tgccagcgtc	agtcccagga	gtccacactg	gtccatcctg	3540
ccacccccaa	ccattttgcac	tacaagggcg	gaggcacccc	gaagaatgaa	aagtacacac	3600
ccatggaatt	caagaccctg	aacaagaata	acttgatccc	tgatgacaga	gccaacttct	3660
acccattgca	gcagaccaat	gtgtacacga	ctacttacta	cccaagcccc	ctgaacaaac	3720
acagcttccg	gcccagggcc	tcacctggac	aacggtgctt	ccccaacagc	tgataccgcc	3780
gtcctgggga	cttggggttc	ttgccttcat	aaggcacaga	gcagatggag	atgggacagt	3840
ggagccagtt	tggttttctc	cctctgcaact	aggccaagaa	cttgcctgct	tgctgtggg	3900
gggtcccatc	cggcttcaga	gagctctggc	tggcattgac	catgggggaa	agggctggtt	3960
tcaggctgac	atatggccgc	aggctccagt	cagcccaggt	ctmtcatggt	tatcttccaa	4020
cccactgtca	cgtcgacact	atgctgccat	gcctgggctg	tggacctact	gggcatattga	4080
ggaaytgag	aatggagatg	gcaagagggc	aggcttttaa	gtttgggttg	gagacaactt	4140
cctgtggccc	ccacaagctg	agtctggcct	tctccagctg	gccccaaaaa	aggcctttgc	4200
tacatcctga	ttatctctga	aagtaatcaa	tcaagtggct	ccagtagctc	tggattttct	4260
gccagggctg	ggccattgtg	gtgctgcccc	agtatgacat	gggaccaagg	ccagcgcagg	4320
ttatccacct	ctgcctggaa	gtctatactc	taccaggggc	atccctctgg	tcagaggcag	4380
tgagtaactg	gaactggagg	ctgacctgtg	cttagaagtc	ctttaatctg	ggctggtaca	4440
ggcctcagcc	ttgccctcaa	tgcacgaaag	gtggcccagg	agagaggatc	aatgccatag	4500
gaggcagaag	tctggcctct	gtgcctctat	ggagactatc	ttccagttgc	tgctcaacag	4560
agttgtttgg	tgagacctgc	ttgggagtct	ctgctggccc	ttcatctgtt	caggaacaca	4620
cacacacaca	cactcacaca	cgcacacaca	atcacaattt	gctacagcaa	caaaaaagac	4680
attgggctgt	ggcattatta	attaaagatg	atatccagtc	aaaaaaaaac	t	4731

<210> 5
 <211> 4703
 <212> DNA
 <213> Homo sapiens

<400> 5						
gcggccgccc	cattcccaga	ccggccgcca	gcccattctg	ttagctcccg	ccgctccgcg	60
ccgcccggga	gtcgggagcc	gcgggggaacc	gggcacctgc	acccgcctct	gggagccagc	120
ttgggtccgg	gttgccatgg	gccctgccag	ggctgtggtc	ggcgcatctg	ggctgcagcg	180
gcgatgggga	cccgggaccc	aggcctggag	aaggagacgg	acgagtgagg	ctgagggacg	240
gaggggacaga	gtgagtgggt	ccagctgggtg	cctggcctgt	gtctcttgga	tgccctgtgg	300
cttcagtcgg	tctcctgttg	cccaccacct	cgctcctggg	ccgcctgata	ccccagccca	360

acagctaagg	tgtggatgga	cagtaggggg	ctggcttctc	tcaactgggtca	ggggctcttct	420
cccctgtctg	cctcccggag	ctaggactgc	agagggggcct	atcatgggtgc	ttgcaggccc	480
cctggctgtc	tcgctgttgc	tgcccagcct	cacactgtctg	gtgtcccacc	tctccagctc	540
ccaggatgtc	tccagtgagc	ccagcagtga	gcagcagctg	tgcgccctta	gcaagcacc	600
caccgtggcc	tttgaagacc	tgacgccgtg	ggctcttaac	ttcacctacc	ctggagccc	660
ggatttctcc	cagctggctt	tggacccctc	cggaaccag	ctcatcgtgg	gagccaggaa	720
ctacctcttc	agactcagcc	ttgccaatgt	ctctcttctt	caggccacag	agtgggcctc	780
cagtgaggac	acgcgccgct	cctgccaaag	caaaggggaag	actgaggagg	agtgtcagaa	840
ctacgtgcga	gtcctgatcg	tcgcccggccg	gaaggtgttc	atgtgtggaa	ccaatgcctt	900
ttcccccatg	tgacccagca	gacaggtggg	gaacctcagc	cggactactg	agaagatcaa	960
tgggtgtggcc	cgctgcccct	atgaccacag	ccacaactcc	acagctgtca	tctcctccca	1020
gggggagctc	tatgcagcca	cggtcatcga	cttctcaggt	cgggaccctg	ccatctaccg	1080
cagcctgggc	agtgggccac	cgcttcgcac	tgcccataat	aactccaagt	ggcttaatga	1140
gccaaacttc	gtggcagcct	atgatattgg	gtggtttgca	tacttcttcc	tgcgggagaa	1200
cgcagtggag	cacgactgtg	gacgcaccgt	gtactctcgc	gtggcccgcg	tgtgcaagaa	1260
tgacgtgggg	ggccgattcc	tgctggagga	cacatggacc	acattcatga	aggcccggct	1320
caactgtctc	cgcccgggcg	aggtcccctt	ctactataac	gagctgcaga	gtgccttcca	1380
cttgccagag	caggaccta	tctatggagt	tttcacaacc	aacgtaaaca	gcatcgcggc	1440
ttctgtctgc	tgcgccctca	acctcagtcg	tatctcccag	gctttcaatg	gcccatttccg	1500
ctaccaggag	aaccccaggg	ctgcctggct	ccccatagcc	aaccccatcc	ccaatttcca	1560
gtgtggcacc	ctgcctgaga	ccggtcccaa	cgagaacctg	acggagcgca	gctctcagga	1620
cgcgcagcgc	ctcttcttga	tgagcgaggg	cggtgacccg	gtgacaccgc	agcctgtgtg	1680
caccagggac	agcgtgcgct	tctcacacct	cggtgtggac	ctggtgcagg	ctaaagacac	1740
gctctaccat	gtactctaca	ttggcaccga	gtcgggcacc	atcctgaagg	cgctgtccac	1800
ggcgagccgc	agcctccacg	gctgctacct	ggaggagctg	cacgtgtctg	ccccggggcg	1860
ccgcgagccc	ctgcgcagcc	tgcgcatcct	gcacagcgcc	cgcgcgctct	tcgtggggct	1920
gagagacggc	gtcctgcggg	tcccactgga	gaggtgcgcc	gcctaccgca	gccagggggc	1980
atgcctgggg	gcccgggacc	cgtactgtgg	ctgggacggg	aagcagcaac	gttgacgac	2040
actcgaggac	agctccaaca	tgagcctctg	gaccagaac	atcaccgcct	gtcctgtgcg	2100
gaatgtgaca	cgggatgggg	gcttcggccc	atggtcacca	tggcaaccat	gtgagcactt	2160
ggatggggac	aactcaggct	cttgccgtgtg	tcgagctcga	tccgtgtgatt	ccccctgacc	2220
ccgctgtggg	ggccttgact	gcctggggcc	agccatccac	atcgccaact	gctccaggaa	2280
tggggcggtg	accccggtgt	catcggtggc	gctgtgcagc	acgtcctgtg	gcatcggtct	2340
ccagggtccgc	cagcgaagtt	gcagcaacct	tgctccccgc	cacggggggc	gcatctgcgt	2400
gggcaagagc	cgggaggaac	ggttctgtaa	tgagaacacg	ccttgcccgcg	tgcccattctt	2460
ctgggcttcc	tggggctcct	ggagcaagtg	cagcagcaac	tgtggagggg	gcatgcagtc	2520
gcggcgctcg	gcctgcgaga	acggcaactc	ctgcctgggc	tgccggcggtg	agttcaagac	2580
gtgcaacccc	gagggctgcc	ccgaagtgcg	gcgcaacacc	ccctggacgc	cgtggctgcc	2640
cgtgaacgtg	acgcagggcg	gggcacggga	ggagcagcgg	ttccgcttca	cctgcgcgcg	2700
gccccttgca	gaccgcacg	gcctgcagtt	cggcaggaga	aggaccgaga	cgaggacctg	2760
tcccgcggac	ggctccggct	cctgcgacac	cgacgccctg	gtggaggtcc	tctgcgcag	2820
cgggagcacc	tccccgcaca	cggtgagcgg	gggctggggc	gcctggggcc	cgtggtcgtc	2880
ctgctcccgc	gactgcgagc	tgggcttccg	cgtccgcaag	agaacgtgca	ctaaccggga	2940
gccccgcaac	gggggcctgc	cctgcgtggg	cgatgtgtcc	gagtaccagg	actgcaacct	3000
ccaggcttgc	ccagttcggg	gtgcttgggt	ctgctggacc	tcatggtctc	catgctcagc	3060
ttcctgtggt	gggggtcact	atcaacgcac	ccgttcctgc	accagccccg	caccctcccc	3120
aggtgaggac	atctgtctcg	ggctgcacac	ggaggaggca	ctatgtgcca	cacaggcctg	3180
cccagaaggc	tggtcgccct	ggtctgagtg	gagtaagtgc	actgacgacg	gagcccagag	3240
ccgaagccgg	cactgtgagg	agtcctccc	agggctcagc	gcmtgtgctg	gaaacagcag	3300
ccagagccgc	ccctgcccc	acagcgagat	tcscgtcatc	ctgccagcct	ccagcatgga	3360
ggaggccacc	gactgtgcag	ggttcaatct	catccacttg	gtggccacgg	gcatctcctg	3420
cttcttgggc	tctgggctcc	tgaccctagc	agtgtacctg	tcttgccagc	actgccagcg	3480
tcagtcctag	gagtcacac	tggtccatcc	tgccaccccc	aaccatttgc	actacaaggg	3540
cggaggcacc	ccgaagaatg	aaaagtacac	acccatggaa	ttcaagaccc	tgaacaagaa	3600
taacttgatc	cctgatgaca	gagccaactt	ctacccattg	cagcagacca	atgtgtacac	3660
gactacttac	tacccaagcc	ccctgaacaa	acacagcttc	cggcccagag	cctcacctgg	3720
acaacgggtg	ttcccaca	gctgataccg	ccgtcctggg	gacttgggct	tcttgccttc	3780
ataaggcaca	gagcagatgg	agatgggaca	gtggagccag	tttggtttcc	tccctctgca	3840
ctaggccaag	aactgtctgc	cttgccgtgtg	gggggtccca	tccggcttca	gagagctctg	3900
gctggcattg	accatggggg	aaagggtgtg	tttcaggctg	acatatggcc	gcagggtccag	3960
ttcagcccag	gtctmtcatg	gttatcttcc	aacccactgt	cacgctgaca	ctatgctgcc	4020

atgcctgggc	tgtggaccta	ctgggcattt	gaggaaytgg	agaatggaga	tggcaagagg	4080
gcaggctttt	aagtttgggt	tggagacaac	ttcctgtggc	ccccacaagc	tgagtctggc	4140
cttctccagc	tggcccaaaa	aaaggccttt	gctacatcct	gattatctct	gaaagtaatc	4200
aatcaagtgg	ctccagtagc	tctggatttt	ctgccagggc	tgggccattg	tggtgctgcc	4260
ccagtatgac	atgggaccaa	ggccagcgca	ggttatccac	ctctgcctgg	aagtctatac	4320
tctaccaggg	gcattccctct	ggtcagaggc	agtgagtact	gggaactgga	ggctgacctg	4380
tgcttagaag	tcctttaatc	tgggctggta	caggcctcag	ccttgccctc	aatgcacgaa	4440
aggtggccca	ggagagagga	tcaatgccat	aggaggcaga	agtctggcct	ctgtgcctct	4500
atggagacta	tcttccagtt	gctgctcaac	agagttgttg	gctgagacct	gcttgggagt	4560
ctctgctggc	ccttcatctg	ttcaggaaca	cacacacaca	cacactcaca	cacgcacaca	4620
caatcacaat	ttgctacagc	aacaaaaaag	acattgggct	gtggcattat	taattaaaga	4680
tgatatccag	tcaaaaaaaaa	act				4703

<210> 6
 <211> 4405
 <212> DNA
 <213> Homo sapiens

<400> 6

gcggccgccc	cattcccaga	ccggccgcca	gcccactctg	ttagctcccg	ccgctccgcg	60
ccgcccggga	gtcgggagcc	gcggggaacc	gggcacctgc	acccgcctct	gggaggtctt	120
ctcccctgtc	tgcctcccgt	agctaggact	gcagaggggc	ctatcatggg	gcttgcaggc	180
cccctggctg	tctcgctggt	gctgcccagc	ctcacactgc	tggtgtccca	cctctccagc	240
tcccaggatg	tctccagtga	gcccagcagt	gagcagcagc	tgtgcgccct	tagcaagcac	300
cccaccgtgg	cctttgaaga	cctgcagccg	tgggtctcta	acttcacctc	ccctggagcc	360
cgggatttct	cccagctggc	tttggacccc	tccgggaacc	agctcatcgt	gggagccagg	420
aactacctct	tcagactcag	ccttgccaat	gtctctcttc	ttcaggccac	agagtgggcc	480
tccagtgagg	acacgcgcgc	ctcctgccaa	agcaaaggga	agactgagga	ggagtgtcag	540
aactacgtgc	gagtcctgat	cgtcgcccgc	cgaaggtgt	tcatgtgtgg	aaccaatgcc	600
ttttccccc	tgtgcaccag	cagacaggtg	gggaacctca	gccggactac	tgagaagatc	660
aatggtgtgg	cccgtgccc	ctatgaccca	cgccacaact	ccacagctgt	catctcctcc	720
cagggggagc	tctatgcagc	cacggtcctc	gacttctcag	gtcgggaccc	tgccatctac	780
cgcagcctgg	gcagtgggcc	accgcttcgc	actgcccaat	ataactccaa	gtggcttaat	840
gagccaaact	tcgtggcagc	ctatgatatt	gggtgttttg	catacttctt	cctgcgggag	900
aacgcagtgg	agcacgactg	tggacgcacc	gtgtactctc	gcgtggcccg	cgtgtgcaag	960
aatgacgtgg	ggggccgatt	cctgctggag	gacacatgga	ccacattcat	gaaggcccgg	1020
ctcaactgct	cccgcgccgg	cgaggtcccc	ttctactata	acgagctgca	gagtgccctc	1080
cacttgccag	agcaggacct	catctatgga	gttttcacaa	ccaacgtaaa	cagcatcgcg	1140
gcttctgctg	tctgcgcctt	caacctcagt	gctatctccc	aggctttcaa	tggcccattt	1200
cgttaccagg	agaaccccag	ggctgcctgg	ctccccatag	ccaaccccat	cccccaattt	1260
cagtgtggca	ccctgcctga	gacgggtccc	aacgagaacc	tgacggagcg	cagcctgcag	1320
gacgcgcagc	gcctcttctt	gatgagcgag	gccgtgcagc	cggtgacacc	cgagccctgt	1380
gtcaccaggg	acagcgtgcg	cttctcacac	ctcgtggtgg	acctggtgca	ggctaaagac	1440
acgctctacc	atgtactcta	cattggcacc	gagtcgggca	ccatcctgaa	ggcgctgtcc	1500
acggcgagcc	gcagcctcca	cggctgctac	ctggaggagc	tgacgtgct	gccccccggg	1560
cgcgcgcagc	ccctgcgcag	cctgcgcctc	ctgcacagcg	cccgcgcgct	cttcgtgggg	1620
ctgagagacg	gcgtcctgcg	ggtcccactg	gagaggtgcg	ccgcctaccg	cagccagggg	1680
gcatgcctgg	gggcccggga	cccgtactgt	ggctgggacg	ggaagcagca	acgttgacgc	1740
acactcgagg	acagctccaa	catgagcctc	tggaccagga	acatcacgcg	ctgtcctgtg	1800
cggaatgtga	cacgggatgg	gggcttcggc	ccatggtcac	catggcaacc	atgtgagcac	1860
ttggatgggg	acaactcagg	ctcttgccct	tgtcgagctc	gatcctgtga	ttcccctcga	1920
ccccgctgtg	ggggccttga	ctgcctgggg	ccagccatcc	acatcgccaa	ctgctccagg	1980
aatggggcgt	ggaccccgtg	gtcatcgtgg	gcgctgtgca	gcacgtcctg	tggcatcggc	2040
ttccagggtcc	gccagcgaa	ttgcagcaac	cctgctcccc	gccacggggg	ccgcatctgc	2100
gtgggcaaga	gccgggagga	acggttctgt	aatgagaaca	cgcttggccc	ggtgcccctc	2160
ttctgggctt	cctggggctc	ctggagcaag	tgcagcagca	actgtggagg	gggcatgcag	2220
tcgcgcgcgc	gggcctgcga	gaacggcaac	tcctgcctgg	gctgcggcgt	ggagttcaag	2280
acgtgcaacc	ccgagggctg	ccccgaagtg	cggcgcaaca	ccccctggac	gccgtggctg	2340
cccgtgaacg	tgacgcaggg	cggggcacgg	caggagcagc	ggttccgctt	cacctgccgc	2400
gcgccccttg	cagacccgca	cggcctgcag	ttcggcagga	gaaggaccga	gacgaggacc	2460

tgtcccgagg	acggctccgg	ctcctgcgac	accgacgccc	tggtggagggt	cctcctgcgc	2520
agcgggagca	cctccccgca	cacgggtgagc	gggggctggg	cgccctgggg	cccgtgggtcg	2580
tcctgctccc	gggactgcga	gctgggcttc	cgcgctccgca	agagaacgtg	cactaaccgcg	2640
gagccccgca	acggggggcct	gccctgctg	ggcgatgctg	ccgagtacca	ggactgcaac	2700
ccccaggctt	gcccagttcg	gggtgcttg	tcctgctgga	cctcatgggtc	tccatgctca	2760
gcttcctgtg	gtgggggtca	ctatcaacgc	acccgttcct	gcaccagccc	cgcaccctcc	2820
ccagggtgag	acatctgtct	cgggctgcac	acggaggagg	cactatgtgc	cacacaggcc	2880
tgcccagaag	gctggctgcc	ctggctctgag	tggagtaagt	gactgacga	cggagcccag	2940
agccgaagcc	ggcactgtga	ggagctcctc	ccagggtcca	gcgcmgtgc	tggaaacagc	3000
agccagagcc	gcccctgccc	ctacagcgag	attcscgtca	tcctgccagc	ctccagcatg	3060
gaggaggcca	ccgactgtgc	agggttcaat	ctcatccact	tggtggccac	gggcatctcc	3120
tgtctcttgg	gctctgggtc	cctgacccta	gcagtgtacc	tgtcttgcca	gcactggcag	3180
cgtcagctccc	aggagtccac	actggtccat	cctgccaccc	ccaaccattt	gcactacaag	3240
ggcggaggca	ccccgaagaa	tgaaaagtac	acacccatgg	aattcaagac	cctgaacaag	3300
aataacttga	tcctgatga	cagagccaac	ttctacccat	tgacgagac	caatgtgtac	3360
acgactactt	actacccaag	ccccctgaac	aaacacagct	tcgggcccga	ggcctcacct	3420
ggacaacggt	gcttccccaa	cagctgatac	cgccgtcctg	gggacttggtg	cttcttgcc	3480
tcataaggca	cagagcagat	ggagatggga	cagtggagcc	agtttggttt	tctccctctg	3540
cactaggcca	agaacttgct	gccttgccctg	tgggggggtcc	catccggctt	cagagagctc	3600
tggttgcat	tgaccatggg	ggaaagggtc	ggtttcagge	tgacatatgg	ccgcagggtcc	3660
agttcagccc	aggtctmtca	tgggttatctt	ccaacccact	gtcacgctga	cactatgtcg	3720
ccatgcctgg	gctgtggacc	tactgggcat	ttgaggaayt	ggagaatgga	gatggcaaga	3780
ggcgaggctt	ttaaagtttg	gttgagaca	acttctgtg	gccccacaa	gctgagtctg	3840
gccttctcca	gctggcccca	aaaaaggcct	ttgctacatc	ctgattatct	ctgaaagtaa	3900
tcaatcaagt	ggctccagta	gctctggatt	ttctgccagg	gctgggcat	tgtggtgctg	3960
ccccagtatg	acatgggacc	aaggccagcg	caggttatcc	acctctgcct	ggaagtctat	4020
actctaccca	gggcatccct	ctggtcagag	gcagtgagta	ctgggaactg	gaggctgacc	4080
tgtgcttaga	agtcctttaa	tctgggctgg	tacaggcctc	agccttgccc	tcaatgcacg	4140
aaaggtggcc	caggagagag	gatcaatgcc	ataggaggca	gaagtctggc	ctctgtgcct	4200
ctatggagac	tatcttccag	ttgctgctca	acagagtgtg	tggtgagac	ctgcttgga	4260
gtctctgctg	gccccttcac	tgttcaggaa	cacacactca	cacacgaca	cacacgaca	4320
cacaatcaca	atgtgctaca	gcaacaaaaa	agacattggg	ctgtggcatt	attaattaaa	4380
gatgatatcc	agtcaaaaaa	aaact				4405

<210> 7

<211> 3938

<212> DNA

<213> Homo sapiens

<400> 7

gcggccgccc	cattcccaga	ccggccgcca	gcccattctg	ttagctcccc	ccgctccgcg	60
ccgcccggga	gtcgggagcc	gcggggaacc	gggcacctgc	acccgcctct	gggagtgagt	120
ggttccagct	ggtgcctggc	ctgtgtctct	tggatgccct	gtggcttcag	tccgtctcct	180
gttgcccacc	acctcgtccc	tgggcccgcct	gataccccag	cccaacagct	aagggtgtgga	240
tggacagtag	ggggctggct	tctctcactg	gtcaggggtc	ttctcccctg	tctgcctccc	300
ggagctagga	ctgcagaggg	gcctatcatg	gtgcttgagc	gccccctggc	tgtctcgtcg	360
ttgctgcccc	gcctcacact	gctggtgtcc	cacctctcca	gctcccagga	tgtctccagt	420
gagcccagca	gtgagcagca	gctgtgcgcc	cttagcaagc	accccaccgt	ggcctttgaa	480
gacctgcagc	cgtgggtctc	taacttcacc	taccctggag	cccgggattt	ctcccagctg	540
gctttggacc	cctccgggaa	ccagctcatc	gtgggagcca	ggaactacct	cttcagactc	600
agccttgcca	atgtctctct	tcttcaggcc	acagagtggg	cctccagtga	ggacacgcgc	660
cgctcctgcc	aaagcaaagg	gaagactgag	gaggagtgtc	agaactacgt	gcgagtcctg	720
atcgctgcgg	gccggaaggt	gttcatgtgt	ggaaccaatg	ccttttcccc	catgtgcacc	780
agcagacagg	tggggaacct	cagccggact	actgagaaga	tcaatggtgt	ggcccgtgc	840
ccctatgacc	cacgccacaa	ctccacagct	gctcatctct	cccaggggga	gctctatgca	900
gccacgggtc	tcgacttctc	aggtcgggac	cctgccatct	accgcagcct	gggcagtggtg	960
ccaccgcttc	gcactgcccc	atataactcc	aagtggctta	atgagccaaa	cttcgtggca	1020
gcctatgata	ttgggtctgtt	tgcatacttc	ttcctgcggg	agaacgcagt	ggagcacgac	1080
tgtggacgca	ccgtgtactc	tcgctgggcc	cgctgtgca	agaatgacgt	ggggggccga	1140
ttcctgctgg	aggacacatg	gaccacattc	atgaaggccc	ggctcaactg	ctcccgcggc	1200

ggcgaggtcc	ccttctacta	taacgagctg	cagagtgcct	tccacttgcc	agagcaggac	1260
ctcatctatg	gagttttcac	aaccaacgta	aacagcatcg	cggcttctgc	tgtctgcgcc	1320
ttcaacctca	gtgctatctc	ccaggctttc	aatggcccat	ttcgctacca	ggagaacccc	1380
agggctgcct	ggctccccat	agccaacccc	atccccaatt	tccagtgtgg	caccctgcct	1440
gagaccggtc	ccaacgagaa	cctgacggag	cgcagcctgc	aggacgcgca	gcgcctcttc	1500
ctgatgagcg	aggccgtgca	gccgggtgaca	cccagaccct	gtgtcaccca	ggacagcgtg	1560
cgcttctcac	acctcgtggt	ggacctggtg	caggctaaag	acacgctcta	ccatgtactc	1620
tacattggca	ccgagtcggg	caccatcctg	aaggcgtgtg	ccacggcgag	ccgcagcctc	1680
cacggctgct	acctggagga	gctgcacgtg	ctgccccccg	ggcgccgcga	gcccctgcgc	1740
agcctgcgca	tcctgcacag	cgcccgcgcg	ctcttcgtgg	ggctgagaga	cggcgtcctg	1800
cgggtcccac	tggagaggtg	cgccgcctac	cgcagccagg	gggcatgcct	gggggccccg	1860
gaccgcgtact	gtggctggga	cgggaagcag	caacgttgca	gcacactcga	ggacagctcc	1920
aacatgagcc	tctggaccca	gaacatcacc	gcctgtcctg	tgcggaatgt	gacacgggat	1980
gggggttcct	gcccattggtc	acccatggcaa	ccatgtgagc	acttggatgg	ggacaactca	2040
ggctcttgcc	tgtgtcgagc	tcgatcctgt	gattccccctc	gaccccgctg	tgggggcctt	2100
gactgcctgg	ggccagccat	ccacatcgcc	aactgctcca	ggaatggggc	gtggacccccg	2160
tggatcatcg	gggcgctgtg	cagcacgtcc	tgtggcatcg	gcttccaggt	ccgccagcga	2220
agttgcagca	accctgctcc	ccgccacggg	ggccgcctct	gcgtgggcaa	gagccgggag	2280
gaacggttct	gtaatgagaa	cacgccttgc	ccggtgccca	tcttctgggc	ttcctggggc	2340
tcctggagca	agtgcagcag	caactgtgga	gggggcatgc	agtcgcggcg	tcgggcctgc	2400
gagaacggca	actcctgcct	gggctgcggc	gtggagttca	agacgtgcaa	ccccgagggc	2460
tgccccgaag	tgcggcgcaa	caccccctgg	acgccgtggc	tgcccgtgaa	cgtgacgcag	2520
ggcggggcac	ggcaggagca	gcggttcgcg	ttcacctgcc	gcgcgccccct	tgcagacccg	2580
cacggcctgc	agttcggcag	gagaaggacc	gagacgagga	cctgtcccgc	ggacggctcc	2640
ggctcctgcg	acaccgacgc	cctggtggag	gtcctcctgc	gcagcgggag	cacctccccg	2700
cacacggtga	gcgggggctg	ggccgcctgg	ggcccgtggt	cgtcctgctc	ccgggactgc	2760
gagctgggct	tcgcgctccg	caagagaacg	tgcactaacc	cggagccccg	caacgggggc	2820
ctgcctgcg	tgggcgatgc	tgccgagtac	caggactgca	acccccaggc	ttgccagatt	2880
cgggggtgctt	ggtcctgctg	gacctcatgg	tctccatgct	cagcttcctg	tgggtgggggt	2940
cactatcaac	gcacccgttc	ctgcaccagc	cccgcaccct	ccccagggtga	ggacatctgt	3000
ctcgggtcgc	acacggagga	ggcactatgt	gccacacagg	cctgcccaga	aggctggtcg	3060
ccctgggtctg	agtggagtaa	gtgcactgac	gacggagccc	agagccgaag	ccggcactgt	3120
gaggagctcc	tcacagggtc	cagcgcmtgt	gctggaaaca	gcagccagag	ccgccccctgc	3180
ccctacagcg	agattcscgt	catcctgcc	gcctccagca	tggaggaggc	caccgactgt	3240
gcagggttca	atctcatcca	cttgggtggc	acgggcatct	cctgcttctt	gggctctggg	3300
ctcctgaccc	tagcagtgtg	cctgtcttgc	cagcactgcc	agcgtcagtc	ccaggagtcc	3360
acactgggtcc	atcctgccac	cccccaacat	ttgcactaca	agggcgggag	caccccgaag	3420
aatgaaaagt	acacacccat	ggaattcaag	accctgaaca	agaataactt	gatccctgat	3480
gacagagcca	acttctaccc	attgcagcag	accaatgcca	gcgcagggtta	tccacctctg	3540
cctggaagtc	tatactctac	ccagggcatc	cctctggtca	gaggcagtga	gtactgggaa	3600
ctggaggctg	acctgtgctt	agaagtcctt	taatctgggc	tggtagaggc	ctcagccttg	3660
ccctcaatgc	acgaaagggtg	gcccaggaga	gaggatcaat	gccataggag	gcagaagtct	3720
ggcctctgtg	cctctatgga	gactatcttc	cagttgctgc	tcaacagagt	tgttggtgta	3780
gacctgcttg	ggagtctctg	ctggcccttc	atctgttcag	gaacacacac	acacacacac	3840
tcacacacgc	acacacaatc	acaatttgct	acagcaacaa	aaaagacatt	gggctgtggc	3900
attattaatt	aaagatgata	tccagtcaaa	aaaaaact			3938

<210> 8
 <211> 1095
 <212> PRT
 <213> Homo sapiens

<400> 8
 Met Val Leu Ala Gly Pro Leu Ala Val Ser Leu Leu Leu Pro Ser Leu
 1 5 10 15
 Thr Leu Leu Val Ser His Leu Ser Ser Ser Gln Asp Val Ser Ser Glu
 20 25 30

Pro Ser Ser Glu Gln Gln Leu Cys Ala Leu Ser Lys His Pro Thr Val

35					40					45					
Ala 50	Phe	Glu	Asp	Leu	Gln	Pro 55	Trp	Val	Ser	Asn	Phe 60	Thr	Tyr	Pro	Gly
Ala 65	Arg	Asp	Phe	Ser	Gln 70	Leu	Ala	Leu	Asp	Pro 75	Ser	Gly	Asn	Gln	Leu 80
Ile	Val	Gly	Ala	Arg 85	Asn	Tyr	Leu	Phe	Arg 90	Leu	Ser	Leu	Ala	Asn 95	Val
Ser	Leu	Leu	Gln 100	Ala	Thr	Glu	Trp	Ala 105	Ser	Ser	Glu	Asp	Thr 110	Arg	Arg
Ser	Cys	Gln 115	Ser	Lys	Gly	Lys	Thr 120	Glu	Glu	Glu	Cys	Gln 125	Asn	Tyr	Val
Arg 130	Val	Leu	Ile	Val	Ala	Gly 135	Arg	Lys	Val	Phe	Met 140	Cys	Gly	Thr	Asn
Ala 145	Phe	Ser	Pro	Met	Cys 150	Thr	Ser	Arg	Gln	Val 155	Gly	Asn	Leu	Ser	Arg 160
Thr	Thr	Glu	Lys	Ile 165	Asn	Gly	Val	Ala	Arg 170	Cys	Pro	Tyr	Asp	Pro 175	Arg
His	Asn	Ser	Thr 180	Ala	Val	Ile	Ser	Ser 185	Gln	Gly	Glu	Leu	Tyr 190	Ala	Ala
Thr	Val	Ile 195	Asp	Phe	Ser	Gly	Arg 200	Asp	Pro	Ala	Ile	Tyr 205	Arg	Ser	Leu
Gly	Ser 210	Gly	Pro	Pro	Leu	Arg 215	Thr	Ala	Gln	Tyr	Asn 220	Ser	Lys	Trp	Leu
Asn 225	Glu	Pro	Asn	Phe	Val 230	Ala	Ala	Tyr	Asp	Ile 235	Gly	Leu	Phe	Ala	Tyr 240
Phe	Phe	Leu	Arg	Glu 245	Asn	Ala	Val	Glu	His 250	Asp	Cys	Gly	Arg	Thr 255	Val
Tyr	Ser	Arg	Val 260	Ala	Arg	Val	Cys	Lys 265	Asn	Asp	Val	Gly	Gly 270	Arg	Phe
Leu	Leu	Glu 275	Asp	Thr	Trp	Thr	Thr 280	Phe	Met	Lys	Ala	Arg 285	Leu	Asn	Cys
Ser	Arg 290	Pro	Gly	Glu	Val	Pro 295	Phe	Tyr	Tyr	Asn	Glu 300	Leu	Gln	Ser	Ala
Phe 305	His	Leu	Pro	Glu	Gln 310	Asp	Leu	Ile	Tyr	Gly 315	Val	Phe	Thr	Thr	Asn 320
Val	Asn	Ser	Ile	Ala 325	Ala	Ser	Ala	Val	Cys 330	Ala	Phe	Asn	Leu	Ser 335	Ala
Ile	Ser	Gln	Ala 340	Phe	Asn	Gly	Pro	Phe 345	Arg	Tyr	Gln	Glu	Asn 350	Pro	Arg
Ala	Ala	Trp 355	Leu	Pro	Ile	Ala	Asn 360	Pro	Ile	Pro	Asn	Phe 365	Gln	Cys	Gly

Thr Leu Pro Glu Thr Gly Pro Asn Glu Asn Leu Thr Glu Arg Ser Leu
 370 375 380
 Gln Asp Ala Gln Arg Leu Phe Leu Met Ser Glu Ala Val Gln Pro Val
 385 390 395 400
 Thr Pro Glu Pro Cys Val Thr Gln Asp Ser Val Arg Phe Ser His Leu
 405 410 415
 Val Val Asp Leu Val Gln Ala Lys Asp Thr Leu Tyr His Val Leu Tyr
 420 425 430
 Ile Gly Thr Glu Ser Gly Thr Ile Leu Lys Ala Leu Ser Thr Ala Ser
 435 440 445
 Arg Ser Leu His Gly Cys Tyr Leu Glu Glu Leu His Val Leu Pro Pro
 450 455 460
 Gly Arg Arg Glu Pro Leu Arg Ser Leu Arg Ile Leu His Ser Ala Arg
 465 470 475 480
 Ala Leu Phe Val Gly Leu Arg Asp Gly Val Leu Arg Val Pro Leu Glu
 485 490 495
 Arg Cys Ala Ala Tyr Arg Ser Gln Gly Ala Cys Leu Gly Ala Arg Asp
 500 505 510
 Pro Tyr Cys Gly Trp Asp Gly Lys Gln Gln Arg Cys Ser Thr Leu Glu
 515 520 525
 Asp Ser Ser Asn Met Ser Leu Trp Thr Gln Asn Ile Thr Ala Cys Pro
 530 535 540
 Val Arg Asn Val Thr Arg Asp Gly Gly Phe Gly Pro Trp Ser Pro Trp
 545 550 555 560
 Gln Pro Cys Glu His Leu Asp Gly Asp Asn Ser Gly Ser Cys Leu Cys
 565 570 575
 Arg Ala Arg Ser Cys Asp Ser Pro Arg Pro Arg Cys Gly Gly Leu Asp
 580 585 590
 Cys Leu Gly Pro Ala Ile His Ile Ala Asn Cys Ser Arg Asn Gly Gly
 595 600 605
 Arg Gly Pro Arg Gly Ala Ser Trp Ala Ala Val Gln Ala Arg Pro Val
 610 615 620
 Ala Ser Gly Phe Gln Val Arg Gln Arg Ser Cys Ser Asn Pro Ala Pro
 625 630 635 640
 Arg His Gly Gly Arg Ile Cys Val Gly Lys Ser Arg Glu Glu Arg Phe
 645 650 655
 Cys Asn Glu Asn Thr Pro Cys Pro Val Pro Ile Phe Trp Ala Ser Trp
 660 665 670
 Gly Ser Trp Ser Lys Cys Ser Ser Asn Cys Gly Gly Gly Met Gln Ser
 675 680 685

Arg Arg Arg Ala Cys Glu Asn Gly Asn Ser Cys Leu Gly Cys Gly Val
 690 695 700
 Glu Phe Lys Thr Cys Asn Pro Glu Gly Cys Pro Glu Val Arg Arg Asn
 705 710 715 720
 Thr Pro Trp Thr Pro Trp Leu Pro Val Asn Val Thr Gln Gly Gly Ala
 725 730 735
 Arg Gln Glu Gln Arg Phe Arg Phe Thr Cys Arg Ala Pro Leu Ala Asp
 740 745 750
 Pro His Gly Leu Gln Phe Gly Arg Arg Arg Thr Glu Thr Arg Thr Cys
 755 760 765
 Pro Ala Asp Gly Ser Gly Ser Cys Asp Thr Asp Ala Leu Val Glu Val
 770 775 780
 Leu Leu Arg Ser Gly Ser Thr Ser Pro His Thr Val Ser Gly Gly Trp
 785 790 795 800
 Ala Ala Trp Gly Pro Trp Ser Ser Cys Ser Arg Asp Cys Glu Leu Gly
 805 810 815
 Phe Arg Val Arg Lys Arg Thr Cys Thr Asn Pro Glu Pro Arg Asn Gly
 820 825 830
 Gly Leu Pro Cys Val Gly Asp Ala Ala Glu Tyr Gln Asp Cys Asn Pro
 835 840 845
 Gln Ala Cys Pro Val Arg Gly Ala Trp Ser Cys Trp Thr Ser Trp Ser
 850 855 860
 Pro Cys Ser Ala Ser Cys Gly Gly Gly His Tyr Gln Arg Thr Arg Ser
 865 870 875 880
 Cys Thr Ser Pro Ala Pro Ser Pro Gly Glu Asp Ile Cys Leu Gly Leu
 885 890 895
 His Thr Glu Glu Ala Leu Cys Ala Thr Gln Ala Cys Pro Glu Gly Trp
 900 905 910
 Ser Pro Trp Ser Glu Trp Ser Lys Cys Thr Asp Asp Gly Ala Gln Ser
 915 920 925
 Arg Ser Arg His Cys Glu Glu Leu Leu Pro Gly Ser Ser Ala Cys Ala
 930 935 940
 Gly Asn Ser Ser Gln Ser Arg Pro Cys Pro Tyr Ser Glu Ile Arg Val
 945 950 955 960
 Ile Leu Pro Ala Ser Ser Met Glu Glu Ala Thr Asp Cys Ala Gly Phe
 965 970 975
 Asn Leu Ile His Leu Val Ala Thr Gly Ile Ser Cys Phe Leu Gly Ser
 980 985 990
 Gly Leu Leu Thr Leu Ala Val Tyr Leu Ser Cys Gln His Cys Gln Arg
 995 1000 1005
 Gln Ser Gln Glu Ser Thr Leu Val His Pro Ala Thr Pro Asn His

1010	1015	1020
Leu His Tyr Lys Gly Gly Gly Thr Pro Lys Asn Glu Lys Tyr Thr 1025 1030 1035		
Pro Met Glu Phe Lys Thr Leu Asn Lys Asn Asn Leu Ile Pro Asp 1040 1045 1050		
Asp Arg Ala Asn Phe Tyr Pro Leu Gln Gln Thr Asn Val Tyr Thr 1055 1060 1065		
Thr Thr Tyr Tyr Pro Ser Pro Leu Asn Lys His Ser Phe Arg Pro 1070 1075 1080		
Glu Ala Ser Pro Gly Gln Arg Cys Phe Pro Asn Ser 1085 1090 1095		
<210> 9		
<211> 1248		
<212> PRT		
<213> Homo sapiens		
<400> 9		
Arg Pro Pro His Ser Gln Thr Gly Arg Gln Pro Ile Trp Leu Ala Pro 1 5 10 15		
Ala Ala Pro Arg Arg Pro Gly Val Gly Ser Arg Gly Glu Pro Gly Thr 20 25 30		
Cys Thr Arg Leu Trp Glu Pro Ala Trp Val Arg Val Ala Leu Gly Pro 35 40 45		
Ala Arg Ala Val Val Gly Ala Ser Gly Leu Gln Arg Arg Trp Gly Pro 50 55 60		
Gly Thr Gln Ala Trp Arg Arg Arg Arg Thr Ser Glu Ala Glu Gly Arg 65 70 75 80		
Arg Asp Arg Val Ser Gly Ser Ser Trp Cys Leu Ala Cys Val Ser Trp 85 90 95		
Met Pro Cys Gly Phe Ser Pro Ser Pro Val Ala His His Leu Val Pro 100 105 110		
Gly Pro Pro Asp Thr Pro Ala Gln Gln Leu Arg Cys Gly Trp Thr Val 115 120 125		
Gly Gly Trp Leu Leu Ser Leu Val Arg Gly Leu Leu Pro Cys Leu Pro 130 135 140		
Pro Gly Ala Arg Thr Ala Glu Gly Pro Ile Met Val Leu Ala Gly Pro 145 150 155 160		
Leu Ala Val Ser Leu Leu Leu Pro Ser Leu Thr Leu Leu Val Ser His 165 170 175		
Leu Ser Ser Ser Gln Asp Val Ser Ser Glu Pro Ser Ser Glu Gln Gln 180 185 190		
Leu Cys Ala Leu Ser Lys His Pro Thr Val Ala Phe Glu Asp Leu Gln		

195					200					205					
Pro	Trp	Val	Ser	Asn	Phe	Thr	Tyr	Pro	Gly	Ala	Arg	Asp	Phe	Ser	Gln
210						215					220				
Leu	Ala	Leu	Asp	Pro	Ser	Gly	Asn	Gln	Leu	Ile	Val	Gly	Ala	Arg	Asn
225					230					235					240
Tyr	Leu	Phe	Arg	Leu	Ser	Leu	Ala	Asn	Val	Ser	Leu	Leu	Gln	Ala	Thr
				245					250					255	
Glu	Trp	Ala	Ser	Ser	Glu	Asp	Thr	Arg	Arg	Ser	Cys	Gln	Ser	Lys	Gly
			260					265					270		
Lys	Thr	Glu	Glu	Glu	Cys	Gln	Asn	Tyr	Val	Arg	Val	Leu	Ile	Val	Ala
		275					280					285			
Gly	Arg	Lys	Val	Phe	Met	Cys	Gly	Thr	Asn	Ala	Phe	Ser	Pro	Met	Cys
	290					295					300				
Thr	Ser	Arg	Gln	Val	Gly	Asn	Leu	Ser	Arg	Thr	Thr	Glu	Lys	Ile	Asn
305					310					315					320
Gly	Val	Ala	Arg	Cys	Pro	Tyr	Asp	Pro	Arg	His	Asn	Ser	Thr	Ala	Val
				325					330					335	
Ile	Ser	Ser	Gln	Gly	Glu	Leu	Tyr	Ala	Ala	Thr	Val	Ile	Asp	Phe	Ser
			340					345					350		
Gly	Arg	Asp	Pro	Ala	Ile	Tyr	Arg	Ser	Leu	Gly	Ser	Gly	Pro	Pro	Leu
		355					360					365			
Arg	Thr	Ala	Gln	Tyr	Asn	Ser	Lys	Trp	Leu	Asn	Glu	Pro	Asn	Phe	Val
	370					375					380				
Ala	Ala	Tyr	Asp	Ile	Gly	Leu	Phe	Ala	Tyr	Phe	Phe	Leu	Arg	Glu	Asn
385					390					395					400
Ala	Val	Glu	His	Asp	Cys	Gly	Arg	Thr	Val	Tyr	Ser	Arg	Val	Ala	Arg
			405						410					415	
Val	Cys	Lys	Asn	Asp	Val	Gly	Gly	Arg	Phe	Leu	Leu	Glu	Asp	Thr	Trp
			420					425					430		
Thr	Thr	Phe	Met	Lys	Ala	Arg	Leu	Asn	Cys	Ser	Arg	Pro	Gly	Glu	Val
		435					440					445			
Pro	Phe	Tyr	Tyr	Asn	Glu	Leu	Gln	Ser	Ala	Phe	His	Leu	Pro	Glu	Gln
	450					455					460				
Asp	Leu	Ile	Tyr	Gly	Val	Phe	Thr	Thr	Asn	Val	Asn	Ser	Ile	Ala	Ala
465					470					475					480
Ser	Ala	Val	Cys	Ala	Phe	Asn	Leu	Ser	Ala	Ile	Ser	Gln	Ala	Phe	Asn
			485					490					495		
Gly	Pro	Phe	Arg	Tyr	Gln	Glu	Asn	Pro	Arg	Ala	Ala	Trp	Leu	Pro	Ile
			500					505					510		
Ala	Asn	Pro	Ile	Pro	Asn	Phe	Gln	Cys	Gly	Thr	Leu	Pro	Glu	Thr	Gly
		515					520					525			

Pro Asn Glu Asn Leu Thr Glu Arg Ser Leu Gln Asp Ala Gln Arg Leu
 530 535 540
 Phe Leu Met Ser Glu Ala Val Gln Pro Val Thr Pro Glu Pro Cys Val
 545 550 555 560
 Thr Gln Asp Ser Val Arg Phe Ser His Leu Val Val Asp Leu Val Gln
 565 570 575
 Ala Lys Asp Thr Leu Tyr His Val Leu Tyr Ile Gly Thr Glu Ser Gly
 580 585 590
 Thr Ile Leu Lys Ala Leu Ser Thr Ala Ser Arg Ser Leu His Gly Cys
 595 600 605
 Tyr Leu Glu Glu Leu His Val Leu Pro Pro Gly Arg Arg Glu Pro Leu
 610 615 620
 Arg Ser Leu Arg Ile Leu His Ser Ala Arg Ala Leu Phe Val Gly Leu
 625 630 635 640
 Arg Asp Gly Val Leu Arg Val Pro Leu Glu Arg Cys Ala Ala Tyr Arg
 645 650 655
 Ser Gln Gly Ala Cys Leu Gly Ala Arg Asp Pro Tyr Cys Gly Trp Asp
 660 665 670
 Gly Lys Gln Gln Arg Cys Ser Thr Leu Glu Asp Ser Ser Asn Met Ser
 675 680 685
 Leu Trp Thr Gln Asn Ile Thr Ala Cys Pro Val Arg Asn Val Thr Arg
 690 695 700
 Asp Gly Gly Phe Gly Pro Trp Ser Pro Trp Gln Pro Cys Glu His Leu
 705 710 715 720
 Asp Gly Asp Asn Ser Gly Ser Cys Leu Cys Arg Ala Arg Ser Cys Asp
 725 730 735
 Ser Pro Arg Pro Arg Cys Gly Gly Leu Asp Cys Leu Gly Pro Ala Ile
 740 745 750
 His Ile Ala Asn Cys Ser Arg Asn Gly Ala Val Asp Pro Val Val Ile
 755 760 765
 Val Gly Arg Cys Ala Ala Thr Ser Cys Gly Ile Gly Phe Gln Val Arg
 770 775 780
 Gln Arg Ser Cys Ser Asn Pro Ala Pro Arg His Gly Gly Arg Ile Cys
 785 790 795 800
 Val Gly Lys Ser Arg Glu Glu Arg Phe Cys Asn Glu Asn Thr Pro Cys
 805 810 815
 Pro Val Pro Ile Phe Trp Ala Ser Trp Gly Ser Trp Ser Lys Cys Ser
 820 825 830
 Ser Asn Cys Gly Gly Gly Met Gln Ser Arg Arg Arg Ala Cys Glu Asn
 835 840 845

Gly Asn Ser Cys Leu Gly Cys Gly Val Glu Phe Lys Thr Cys Asn Pro
 850 855 860
 Glu Gly Cys Pro Glu Val Arg Arg Asn Thr Pro Trp Thr Pro Trp Leu
 865 870 875 880
 Pro Val Asn Val Thr Gln Gly Gly Ala Arg Gln Glu Gln Arg Phe Arg
 885 890 895
 Phe Thr Cys Arg Ala Pro Leu Ala Asp Pro His Gly Leu Gln Phe Gly
 900 905 910
 Arg Arg Arg Thr Glu Thr Arg Thr Cys Pro Ala Asp Gly Ser Gly Ser
 915 920 925
 Cys Asp Thr Asp Ala Leu Val Glu Val Leu Leu Arg Ser Gly Ser Thr
 930 935 940
 Ser Pro His Thr Val Ser Gly Gly Trp Ala Ala Trp Gly Pro Trp Ser
 945 950 955 960
 Ser Cys Ser Arg Asp Cys Glu Leu Gly Phe Arg Val Arg Lys Arg Thr
 965 970 975
 Cys Thr Asn Pro Glu Pro Arg Asn Gly Gly Leu Pro Cys Val Gly Asp
 980 985 990
 Ala Ala Glu Tyr Gln Asp Cys Asn Pro Gln Ala Cys Pro Val Arg Gly
 995 1000 1005
 Ala Trp Ser Cys Trp Thr Ser Trp Ser Pro Cys Ser Ala Ser Cys
 1010 1015 1020
 Gly Gly Gly His Tyr Gln Arg Thr Arg Ser Cys Thr Ser Pro Ala
 1025 1030 1035
 Pro Ser Pro Gly Glu Asp Ile Cys Leu Gly Leu His Thr Glu Glu
 1040 1045 1050
 Ala Leu Cys Ala Thr Gln Ala Cys Pro Glu Gly Trp Ser Pro Trp
 1055 1060 1065
 Ser Glu Trp Ser Lys Cys Thr Asp Asp Gly Ala Gln Ser Arg Ser
 1070 1075 1080
 Arg His Cys Glu Glu Leu Leu Pro Gly Ser Ser Ala Cys Ala Gly
 1085 1090 1095
 Asn Ser Ser Gln Ser Arg Pro Cys Pro Tyr Ser Glu Ile Arg Val
 1100 1105 1110
 Ile Leu Pro Ala Ser Ser Met Glu Glu Ala Thr Asp Cys Ala Gly
 1115 1120 1125
 Phe Asn Leu Ile His Leu Val Ala Thr Gly Ile Ser Cys Phe Leu
 1130 1135 1140
 Gly Ser Gly Leu Leu Thr Leu Ala Val Tyr Leu Ser Cys Gln His
 1145 1150 1155
 Cys Gln Arg Gln Ser Gln Glu Ser Thr Leu Val His Pro Ala Thr

1160	1165	1170
Pro Asn His Leu His Tyr Lys Gly Gly Gly Thr Pro Lys Asn Glu		
1175	1180	1185
Lys Tyr Thr Pro Met Glu Phe Lys Thr Leu Asn Lys Asn Asn Leu		
1190	1195	1200
Ile Pro Asp Asp Arg Ala Asn Phe Tyr Pro Leu Gln Gln Thr Asn		
1205	1210	1215
Val Tyr Thr Thr Thr Tyr Tyr Pro Ser Pro Leu Asn Lys His Ser		
1220	1225	1230
Phe Arg Pro Glu Ala Ser Pro Gly Gln Arg Cys Phe Pro Asn Ser		
1235	1240	1245

<210> 10
 <211> 1150
 <212> PRT
 <213> Homo sapiens

<400> 10
 Ala Ala Ala Pro Phe Pro Asp Arg Pro Pro Ala His Leu Val Ser Ser
 1 5 10 15
 Arg Arg Ser Ala Pro Pro Gly Ser Arg Glu Pro Arg Gly Thr Gly His
 20 25 30
 Leu His Pro Pro Leu Gly Gly Leu Leu Pro Cys Leu Pro Pro Gly Ala
 35 40 45
 Arg Thr Ala Glu Gly Pro Ile Met Val Leu Ala Gly Pro Leu Ala Val
 50 55 60
 Ser Leu Leu Leu Pro Ser Leu Thr Leu Leu Val Ser His Leu Ser Ser
 65 70 75 80
 Ser Gln Asp Val Ser Ser Glu Pro Ser Ser Glu Gln Gln Leu Cys Ala
 85 90 95
 Leu Ser Lys His Pro Thr Val Ala Phe Glu Asp Leu Gln Pro Trp Val
 100 105 110
 Ser Asn Phe Thr Tyr Pro Gly Ala Arg Asp Phe Ser Gln Leu Ala Leu
 115 120 125
 Asp Pro Ser Gly Asn Gln Leu Ile Val Gly Ala Arg Asn Tyr Leu Phe
 130 135 140
 Arg Leu Ser Leu Ala Asn Val Ser Leu Leu Gln Ala Thr Glu Trp Ala
 145 150 155 160
 Ser Ser Glu Asp Thr Arg Arg Ser Cys Gln Ser Lys Gly Lys Thr Glu
 165 170 175
 Glu Glu Cys Gln Asn Tyr Val Arg Val Leu Ile Val Ala Gly Arg Lys
 180 185 190
 Val Phe Met Cys Gly Thr Asn Ala Phe Ser Pro Met Cys Thr Ser Arg

195					200					205					
Gln	Val	Gly	Asn	Leu	Ser	Arg	Thr	Thr	Glu	Lys	Ile	Asn	Gly	Val	Ala
210						215					220				
Arg	Cys	Pro	Tyr	Asp	Pro	Arg	His	Asn	Ser	Thr	Ala	Val	Ile	Ser	Ser
225					230					235					240
Gln	Gly	Glu	Leu	Tyr	Ala	Ala	Thr	Val	Ile	Asp	Phe	Ser	Gly	Arg	Asp
				245					250					255	
Pro	Ala	Ile	Tyr	Arg	Ser	Leu	Gly	Ser	Gly	Pro	Pro	Leu	Arg	Thr	Ala
			260					265						270	
Gln	Tyr	Asn	Ser	Lys	Trp	Leu	Asn	Glu	Pro	Asn	Phe	Val	Ala	Ala	Tyr
		275					280							285	
Asp	Ile	Gly	Leu	Phe	Ala	Tyr	Phe	Phe	Leu	Arg	Glu	Asn	Ala	Val	Glu
		290				295					300				
His	Asp	Cys	Gly	Arg	Thr	Val	Tyr	Ser	Arg	Val	Ala	Arg	Val	Cys	Lys
305					310					315					320
Asn	Asp	Val	Gly	Gly	Arg	Phe	Leu	Leu	Glu	Asp	Thr	Trp	Thr	Thr	Phe
				325					330					335	
Met	Lys	Ala	Arg	Leu	Asn	Cys	Ser	Arg	Pro	Gly	Glu	Val	Pro	Phe	Tyr
			340					345					350		
Tyr	Asn	Glu	Leu	Gln	Ser	Ala	Phe	His	Leu	Pro	Glu	Gln	Asp	Leu	Ile
		355					360					365			
Tyr	Gly	Val	Phe	Thr	Thr	Asn	Val	Asn	Ser	Ile	Ala	Ala	Ser	Ala	Val
		370				375					380				
Cys	Ala	Phe	Asn	Leu	Ser	Ala	Ile	Ser	Gln	Ala	Phe	Asn	Gly	Pro	Phe
385					390					395					400
Arg	Tyr	Gln	Glu	Asn	Pro	Arg	Ala	Ala	Trp	Leu	Pro	Ile	Ala	Asn	Pro
				405					410					415	
Ile	Pro	Asn	Phe	Gln	Cys	Gly	Thr	Leu	Pro	Glu	Thr	Gly	Pro	Asn	Glu
			420					425					430		
Asn	Leu	Thr	Glu	Arg	Ser	Leu	Gln	Asp	Ala	Gln	Arg	Leu	Phe	Leu	Met
		435					440						445		
Ser	Glu	Ala	Val	Gln	Pro	Val	Thr	Pro	Glu	Pro	Cys	Val	Thr	Gln	Asp
		450				455					460				
Ser	Val	Arg	Phe	Ser	His	Leu	Val	Val	Asp	Leu	Val	Gln	Ala	Lys	Asp
465					470					475					480
Thr	Leu	Tyr	His	Val	Leu	Tyr	Ile	Gly	Thr	Glu	Ser	Gly	Thr	Ile	Leu
			485						490					495	
Lys	Ala	Leu	Ser	Thr	Ala	Ser	Arg	Ser	Leu	His	Gly	Cys	Tyr	Leu	Glu
			500					505					510		
Glu	Leu	His	Val	Leu	Pro	Pro	Gly	Arg	Arg	Glu	Pro	Leu	Arg	Ser	Leu
		515					520						525		

Arg Ile Leu His Ser Ala Arg Ala Leu Phe Val Gly Leu Arg Asp Gly
 530 535 540
 Val Leu Arg Val Pro Leu Glu Arg Cys Ala Ala Tyr Arg Ser Gln Gly
 545 550 555 560
 Ala Cys Leu Gly Ala Arg Asp Pro Tyr Cys Gly Trp Asp Gly Lys Gln
 565 570 575
 Gln Arg Cys Ser Thr Leu Glu Asp Ser Ser Asn Met Ser Leu Trp Thr
 580 585 590
 Gln Asn Ile Thr Ala Cys Pro Val Arg Asn Val Thr Arg Asp Gly Gly
 595 600 605
 Phe Gly Pro Trp Ser Pro Trp Gln Pro Cys Glu His Leu Asp Gly Asp
 610 615 620
 Asn Ser Gly Ser Cys Leu Cys Arg Ala Arg Ser Cys Asp Ser Pro Arg
 625 630 635 640
 Pro Arg Cys Gly Gly Leu Asp Cys Leu Gly Pro Ala Ile His Ile Ala
 645 650 655
 Asn Cys Ser Arg Asn Gly Ala Val Asp Pro Val Val His Arg Gly Pro
 660 665 670
 Leu Cys Ser His Val Leu Trp His Ala Ala Ser Arg Ser Ala Ser Glu
 675 680 685
 Val Ala Ala Thr Leu Leu Pro Ala Thr Gly Ala Ala Ser Ala Trp Ala
 690 695 700
 Arg Ala Trp Glu Glu Arg Phe Cys Asn Glu Asn Thr Pro Cys Pro Val
 705 710 715 720
 Pro Ile Phe Trp Ala Ser Trp Gly Ser Trp Ser Lys Cys Ser Ser Asn
 725 730 735
 Cys Gly Gly Gly Met Gln Ser Arg Arg Arg Ala Cys Glu Asn Gly Asn
 740 745 750
 Ser Cys Leu Gly Cys Gly Val Glu Phe Lys Thr Cys Asn Pro Glu Gly
 755 760 765
 Cys Pro Glu Val Arg Arg Asn Thr Pro Trp Thr Pro Trp Leu Pro Val
 770 775 780
 Asn Val Thr Gln Gly Gly Ala Arg Gln Glu Gln Arg Phe Arg Phe Thr
 785 790 795 800
 Cys Arg Ala Pro Leu Ala Asp Pro His Gly Leu Gln Phe Gly Arg Arg
 805 810 815
 Arg Thr Glu Thr Arg Thr Cys Pro Ala Asp Gly Ser Gly Ser Cys Asp
 820 825 830
 Thr Asp Ala Leu Val Glu Val Leu Leu Arg Ser Gly Ser Thr Ser Pro
 835 840 845

His Thr Val Ser Gly Gly Trp Ala Ala Trp Gly Pro Trp Ser Ser Cys
 850 855 860
 Ser Arg Asp Cys Glu Leu Gly Phe Arg Val Arg Lys Arg Thr Cys Thr
 865 870 875 880
 Asn Pro Glu Pro Arg Asn Gly Gly Leu Pro Cys Val Gly Asp Ala Ala
 885 890 895
 Glu Tyr Gln Asp Cys Asn Pro Gln Ala Cys Pro Val Arg Gly Ala Trp
 900 905 910
 Ser Cys Trp Thr Ser Trp Ser Pro Cys Ser Ala Ser Cys Gly Gly Gly
 915 920 925
 His Tyr Gln Arg Thr Arg Ser Cys Thr Ser Pro Ala Pro Ser Pro Gly
 930 935 940
 Glu Asp Ile Cys Leu Gly Leu His Thr Glu Glu Ala Leu Cys Ala Thr
 945 950 955 960
 Gln Ala Cys Pro Glu Gly Trp Ser Pro Trp Ser Glu Trp Ser Lys Cys
 965 970 975
 Thr Asp Asp Gly Ala Gln Ser Arg Ser Arg His Cys Glu Glu Leu Leu
 980 985 990
 Pro Gly Ser Ser Ala Cys Ala Gly Asn Ser Ser Gln Ser Arg Pro Cys
 995 1000 1005
 Pro Tyr Ser Glu Ile Arg Val Ile Leu Pro Ala Ser Ser Met Glu
 1010 1015 1020
 Glu Ala Thr Asp Cys Ala Gly Phe Asn Leu Ile His Leu Val Ala
 1025 1030 1035
 Thr Gly Ile Ser Cys Phe Leu Gly Ser Gly Leu Leu Thr Leu Ala
 1040 1045 1050
 Val Tyr Leu Ser Cys Gln His Cys Gln Arg Gln Ser Gln Glu Ser
 1055 1060 1065
 Thr Leu Val His Pro Ala Thr Pro Asn His Leu His Tyr Lys Gly
 1070 1075 1080
 Gly Gly Thr Pro Lys Asn Glu Lys Tyr Thr Pro Met Glu Phe Lys
 1085 1090 1095
 Thr Leu Asn Lys Asn Asn Leu Ile Pro Asp Asp Arg Ala Asn Phe
 1100 1105 1110
 Tyr Pro Leu Gln Gln Thr Asn Val Tyr Thr Thr Thr Tyr Tyr Pro
 1115 1120 1125
 Ser Pro Leu Asn Lys His Ser Phe Arg Pro Glu Ala Ser Pro Gly
 1130 1135 1140
 Gln Arg Cys Phe Pro Asn Ser
 1145 1150

<210> 11
 <211> 1211
 <212> PRT
 <213> Homo sapiens

<400> 11
 Ala Ala Ala Pro Phe Pro Asp Arg Pro Pro Ala His Leu Val Ser Ser
 1 5 10 15
 Arg Arg Ser Ala Pro Pro Gly Ser Arg Glu Pro Arg Gly Thr Gly His
 20 25 30
 Leu His Pro Pro Leu Gly Val Ser Gly Ser Ser Trp Cys Leu Ala Cys
 35 40 45
 Val Ser Trp Met Pro Cys Gly Phe Ser Pro Ser Pro Val Ala His His
 50 55 60
 Leu Val Pro Gly Pro Pro Asp Thr Pro Ala Gln Gln Leu Arg Cys Gly
 65 70 75 80
 Trp Thr Val Gly Gly Trp Leu Leu Ser Leu Val Arg Gly Leu Leu Pro
 85 90 95
 Cys Leu Pro Pro Gly Ala Arg Thr Ala Glu Gly Pro Ile Met Val Leu
 100 105 110
 Ala Gly Pro Leu Ala Val Ser Leu Leu Leu Pro Ser Leu Thr Leu Leu
 115 120 125
 Val Ser His Leu Ser Ser Ser Gln Asp Val Ser Ser Glu Pro Ser Ser
 130 135 140
 Glu Gln Gln Leu Cys Ala Leu Ser Lys His Pro Thr Val Ala Phe Glu
 145 150 155 160
 Asp Leu Gln Pro Trp Val Ser Asn Phe Thr Tyr Pro Gly Ala Arg Asp
 165 170 175
 Phe Ser Gln Leu Ala Leu Asp Pro Ser Gly Asn Gln Leu Ile Val Gly
 180 185 190
 Ala Arg Asn Tyr Leu Phe Arg Leu Ser Leu Ala Asn Val Ser Leu Leu
 195 200 205
 Gln Ala Thr Glu Trp Ala Ser Ser Glu Asp Thr Arg Arg Ser Cys Gln
 210 215 220
 Ser Lys Gly Lys Thr Glu Glu Glu Cys Gln Asn Tyr Val Arg Val Leu
 225 230 235 240
 Ile Val Ala Gly Arg Lys Val Phe Met Cys Gly Thr Asn Ala Phe Ser
 245 250 255
 Pro Met Cys Thr Ser Arg Gln Val Gly Asn Leu Ser Arg Thr Thr Glu
 260 265 270
 Lys Ile Asn Gly Val Ala Arg Cys Pro Tyr Asp Pro Arg His Asn Ser
 275 280 285
 Thr Ala Val Ile Ser Ser Gln Gly Glu Leu Tyr Ala Ala Thr Val Ile

290	295	300
Asp Phe Ser Gly Arg Asp Pro Ala Ile Tyr Arg Ser Leu Gly Ser Gly 305 310 315 320		
Pro Pro Leu Arg Thr Ala Gln Tyr Asn Ser Lys Trp Leu Asn Glu Pro 325 330 335		
Asn Phe Val Ala Ala Tyr Asp Ile Gly Leu Phe Ala Tyr Phe Phe Leu 340 345 350		
Arg Glu Asn Ala Val Glu His Asp Cys Gly Arg Thr Val Tyr Ser Arg 355 360 365		
Val Ala Arg Val Cys Lys Asn Asp Val Gly Gly Arg Phe Leu Leu Glu 370 375 380		
Asp Thr Trp Thr Thr Phe Met Lys Ala Arg Leu Asn Cys Ser Arg Pro 385 390 395 400		
Gly Glu Val Pro Phe Tyr Tyr Asn Glu Leu Gln Ser Ala Phe His Leu 405 410 415		
Pro Glu Gln Asp Leu Ile Tyr Gly Val Phe Thr Thr Asn Val Asn Ser 420 425 430		
Ile Ala Ala Ser Ala Val Cys Ala Phe Asn Leu Ser Ala Ile Ser Gln 435 440 445		
Ala Phe Asn Gly Pro Phe Arg Tyr Gln Glu Asn Pro Arg Ala Ala Trp 450 455 460		
Leu Pro Ile Ala Asn Pro Ile Pro Asn Phe Gln Cys Gly Thr Leu Pro 465 470 475 480		
Glu Thr Gly Pro Asn Glu Asn Leu Thr Glu Arg Ser Leu Gln Asp Ala 485 490 495		
Gln Arg Leu Phe Leu Met Ser Glu Ala Val Gln Pro Val Thr Pro Glu 500 505 510		
Pro Cys Val Thr Gln Asp Ser Val Arg Phe Ser His Leu Val Val Asp 515 520 525		
Leu Val Gln Ala Lys Asp Thr Leu Tyr His Val Leu Tyr Ile Gly Thr 530 535 540		
Glu Ser Gly Thr Ile Leu Lys Ala Leu Ser Thr Ala Ser Arg Ser Leu 545 550 555 560		
His Gly Cys Tyr Leu Glu Glu Leu His Val Leu Pro Pro Gly Arg Arg 565 570 575		
Glu Pro Leu Arg Ser Leu Arg Ile Leu His Ser Ala Arg Ala Leu Phe 580 585 590		
Val Gly Leu Arg Asp Gly Val Leu Arg Val Pro Leu Glu Arg Cys Ala 595 600 605		
Ala Tyr Arg Ser Gln Gly Ala Cys Leu Gly Ala Arg Asp Pro Tyr Cys 610 615 620		

Gly	Trp	Asp	Gly	Lys	Gln	Gln	Arg	Cys	Ser	Thr	Leu	Glu	Asp	Ser	Ser	
625					630					635					640	
Asn	Met	Ser	Leu	Trp	Thr	Gln	Asn	Ile	Thr	Ala	Cys	Pro	Val	Arg	Asn	
			645						650					655		
Val	Thr	Arg	Asp	Gly	Gly	Phe	Gly	Pro	Trp	Ser	Pro	Trp	Gln	Pro	Cys	
			660					665					670			
Glu	His	Leu	Asp	Gly	Asp	Asn	Ser	Gly	Ser	Cys	Leu	Cys	Arg	Ala	Arg	
		675					680					685				
Ser	Cys	Asp	Ser	Pro	Arg	Pro	Arg	Cys	Gly	Gly	Leu	Asp	Cys	Leu	Gly	
	690					695					700					
Pro	Ala	Ile	His	Ile	Ala	Asn	Cys	Ser	Arg	Asn	Gly	Ala	Val	Asp	Pro	
705					710					715					720	
Val	Val	Ile	Val	Gly	Arg	Cys	Ala	Ala	Thr	Ser	Cys	Gly	Ile	Gly	Phe	
			725						730					735		
Gln	Val	Arg	Gln	Arg	Ser	Cys	Ser	Asn	Pro	Ala	Pro	Arg	His	Gly	Gly	
			740					745					750			
Arg	Ile	Cys	Val	Gly	Lys	Ser	Arg	Glu	Glu	Arg	Phe	Cys	Asn	Glu	Asn	
		755					760					765				
Thr	Pro	Cys	Pro	Val	Pro	Ile	Phe	Trp	Ala	Ser	Trp	Gly	Ser	Trp	Ser	
	770					775					780					
Lys	Cys	Ser	Ser	Asn	Cys	Gly	Gly	Gly	Met	Gln	Ser	Arg	Arg	Arg	Ala	
785					790					795					800	
Cys	Glu	Asn	Gly	Asn	Ser	Cys	Leu	Gly	Cys	Gly	Val	Glu	Phe	Lys	Thr	
			805						810					815		
Cys	Asn	Pro	Glu	Gly	Cys	Pro	Glu	Val	Arg	Arg	Asn	Thr	Pro	Trp	Thr	
			820					825					830			
Pro	Trp	Leu	Pro	Val	Asn	Val	Thr	Gln	Gly	Gly	Ala	Arg	Gln	Glu	Gln	
		835					840					845				
Arg	Phe	Arg	Phe	Thr	Cys	Arg	Ala	Pro	Leu	Ala	Asp	Pro	His	Gly	Leu	
	850					855					860					
Gln	Phe	Gly	Arg	Arg	Arg	Thr	Glu	Thr	Arg	Thr	Cys	Pro	Ala	Asp	Gly	
865					870					875					880	
Ser	Gly	Ser	Cys	Asp	Thr	Asp	Ala	Leu	Val	Glu	Val	Leu	Leu	Arg	Ser	
				885					890					895		
Gly	Ser	Thr	Ser	Pro	His	Thr	Val	Ser	Gly	Gly	Trp	Ala	Ala	Trp	Gly	
			900					905					910			
Pro	Trp	Ser	Ser	Cys	Ser	Arg	Asp	Cys	Glu	Leu	Gly	Phe	Arg	Val	Arg	
		915					920					925				
Lys	Arg	Thr	Cys	Thr	Asn	Pro	Glu	Pro	Arg	Asn	Gly	Gly	Leu	Pro	Cys	
	930					935					940					

Val Gly Asp Ala Ala Glu Tyr Gln Asp Cys Asn Pro Gln Ala Cys Pro
 945 950 955 960
 Val Arg Gly Ala Trp Ser Cys Trp Thr Ser Trp Ser Pro Cys Ser Ala
 965 970 975
 Ser Cys Gly Gly Gly His Tyr Gln Arg Thr Arg Ser Cys Thr Ser Pro
 980 985 990
 Ala Pro Ser Pro Gly Glu Asp Ile Cys Leu Gly Leu His Thr Glu Glu
 995 1000 1005
 Ala Leu Cys Ala Thr Gln Ala Cys Pro Glu Gly Trp Ser Pro Trp
 1010 1015 1020
 Ser Glu Trp Ser Lys Cys Thr Asp Asp Gly Ala Gln Ser Arg Ser
 1025 1030 1035
 Arg His Cys Glu Glu Leu Leu Pro Gly Ser Ser Ala Cys Ala Gly
 1040 1045 1050
 Asn Ser Ser Gln Ser Arg Pro Cys Pro Tyr Ser Glu Ile Arg Val
 1055 1060 1065
 Ile Leu Pro Ala Ser Ser Met Glu Glu Ala Thr Asp Cys Ala Gly
 1070 1075 1080
 Phe Asn Leu Ile His Leu Val Ala Thr Gly Ile Ser Cys Phe Leu
 1085 1090 1095
 Gly Ser Gly Leu Leu Thr Leu Ala Val Tyr Leu Ser Cys Gln His
 1100 1105 1110
 Cys Gln Arg Gln Ser Gln Glu Ser Thr Leu Val His Pro Ala Thr
 1115 1120 1125
 Pro Asn His Leu His Tyr Lys Gly Gly Gly Thr Pro Lys Asn Glu
 1130 1135 1140
 Lys Tyr Thr Pro Met Glu Phe Lys Thr Leu Asn Lys Asn Asn Leu
 1145 1150 1155
 Ile Pro Asp Asp Arg Ala Asn Phe Tyr Pro Leu Gln Gln Thr Asn
 1160 1165 1170
 Ala Ser Ala Gly Tyr Pro Pro Leu Pro Gly Ser Leu Tyr Ser Thr
 1175 1180 1185
 Gln Gly Ile Pro Leu Val Arg Gly Ser Glu Tyr Trp Glu Leu Glu
 1190 1195 1200
 Ala Asp Leu Cys Leu Glu Val Leu
 1205 1210

<210> 12
 <211> 1203
 <212> PRT
 <213> Homo sapiens

<400> 12